Computational Models of Attention and Cognitive Control

1. Introduction

The study of attention is central to understanding how information is processed in cognitive systems. Modern cognitive research interprets attention as the capacity to select and enhance limited aspects of currently processed information, while suppressing the remaining aspects. Cognitive scientists interpret attention as a solution to a fundamental computational trade-off that limited agents face in complex environments: on one side, the necessity to focus on as much information as possible in order to be vigilant and opportunistic, on the other side, the necessity to optimize performance by allocating, in a coherent and continuous manner, cognitive resources to the most salient and behaviorally relevant events and actions (Allport, 1989). As such, attention turns out not to be a unitary phenomenon, but instead is present at many stages of cognitive information processing, involves many different brain regions, and relates to almost all psychological processes.

This chapter reviews the existing literature on computational models of attention, with the aim of fleshing out the progress that has been made in elucidating the core mechanisms of attentional modulation and attentional control. The chapter starts with a description of work that focuses on visual selective attention and the computational mechanisms that exist at the site of attentional influence within visual perceptual pathways. Subsequent sections focus on work at the intersection of attention and executive control, which emphasizes the mechanisms by which goal-driven attentional control signals are represented, shaped, and propagated according to the various constraints and dynamics of task processing. In the concluding section, the focus is on the contrast or continuum between attentional control and automaticity, an issue that becomes crystallized when examining the distinctions between, or transitions from, novice to expert cognitive task performance.

It is important to begin with a caveat – this chapter is not intended to be comprehensive or exhaustive in the coverage of computational cognitive modeling work on attention. Instead, the goal is to provide a

road map to the relevant literature, highlighting example models that best reflect the core mechanistic principles that are emerging from recent research or that illustrate new directions in which the field is headed. Moreover, the coverage is admittedly biased toward connectionist or neural network models. The reason for this bias is not only due to the expertise of the authors, but also to an overarching interest in computational models that have the most potential for integrating the large emerging corpus of literature, arising not only out of cognitive behavioral research, but also cognitive neuroscience and animal neurophysiology studies. Traditionally, this approach has been most closely aligned with connectionist/neural network models, although recent trends suggest that this traditional dichotomy between connectionist and symbolic models is beginning to blur (e.g., Anderson et al., 2004).

Nevertheless, it is still our belief that models that make strong attempts to incorporate as many core principles of neural information processing and computation as possible are the ones most likely to explain empirical data regarding attentional phenomena across the widest range of explanatory levels, from single-cell neurophysiology to observable behavior. Although this philosophical bias is reflected throughout the chapter, influential examples are also reviewed from work in symbolic, hybrid, or production system modeling, as well as more abstract mathematical models. Readers interested in learning more about both neurally oriented models of attention and symbolic ones are directed toward the many additional reviews of this literature arising from a variety of different theoretical perspectives and focus (i.e., Anderson et al., 2004; Itti & Koch, 2001; O'Reilly & Munakata, 2000).

2. Visual Attention

When we observe and interact with our environment, the focus of what we are attending to constantly changes. There are a variety of theoretical views regarding why attention selectively focuses on some aspects of the environment and away from others. Perhaps the oldest argument is that our processing capabilities are limited whereas the computational demands of processing visual input are huge. Under this account, the role of attention is to filter this spatiotemporal stream of information to a manageable size (Broadbent, 1958; Mozer, Sitton, & Pashler, 1998). A second theoretical argument regarding selection in visual attention is that not all of the information present in the visual environment is equally relevant at a given point in time. Under this account, the role of attention is to quickly detect, orient, and select the aspects of the visual environment that are most informative or of greatest relevance at the time, so as to produce efficient and optimized perceptual processing and subsequent behavior (Chang et al., 2001; van der Heijden & Bem, 1997). Yet, a third theoretical perspective is that the primary role of visual attention is to solve the binding problem: to produce a coherent interpretation of the visual environment based on integration of visual features into a unified whole. Under this account, selective attention enables visual perceptual processing to be concentrated on a restricted set of visual features to enable these to be correctly bound together into higher-level object representations (Treisman, 1999; Treisman & Gelade, 1980). And, lastly, a more recent theoretical view is the biased competition framework, which postulates that attention should be interpreted primarily as an emergent phenomenon of activation dynamics arising in a system in which inhibitory competition and constraint satisfaction is a ubiquitous component of the network (Desimone & Duncan, 1995; O'Reilly & Munakata, 2000).
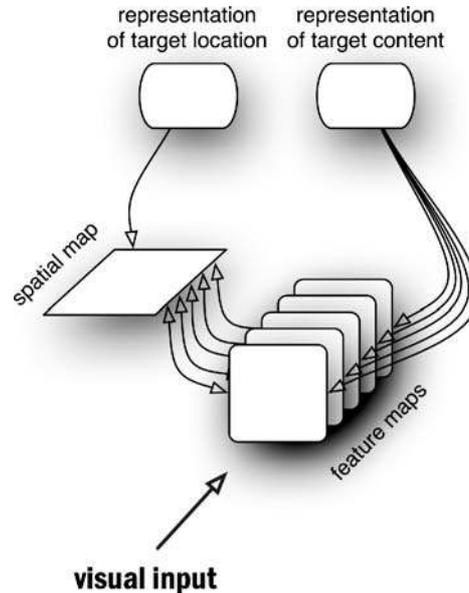
Regardless of the particular theoretical perspective one adopts, there is clearly a consensus among theorists that: (1) attention is a core component of visual perceptual processing; (2) focus of attention is determined by an interaction of bottom-up processes that compute the "importance" of visual stimuli and top-down processes that modulate visual processing according

to goals and intentions; and (3) top-down processes can operate by directing attention either to locations in space or to specific objects or object features in the visual field.

A number of theories of visual attention have been implemented in computational models. The design and scope of the existing computational models of visual attention vary widely and are determined by the problems they are meant to resolve. Some models were built to explicitly test specific aspects of existing theories of visual attention or to account for empirical data that apply to specific experimental domains, such as stimulus filtering, visual search, and perceptual cueing (Cave, 1999; Deco & Zihl, 2001a; Heinke & Humphreys, 2003; Humphreys & Mueller, 1993; Mozer et al., 1998; Phaf, Van der Heijden, & Hudson 1990; Wolfe, 1994). Other models were particularly geared toward explaining the mechanisms underlying particular neuropsychological deficits, such as attentional neglect (Cohen et al., 1994; Deco & Rolls, 2002; Heinke & Humphreys, 2003). Still other models were designed primarily from a neurophysiological perspective to account for basic aspects of early visual processing and identification of salient locations in visual field (Itti & Koch, 2000; Koch & Ullman, 1985; Lee et al., 1999; Parkhurst, Law, & Niebur, 2002) or to understand the core neurobiological mechanisms involved in attention (Braun, Koch, & Davis, 2001; Hamker, 2003). Finally, some models were built primarily from a machine learning rather than cognitive, neuropsychological, or neurobiological perspective, such as advancing the development of computer vision systems (Tsotsos et al., 1995). Rather than detailing each one of these models, the following subsections focus on what might be considered a "consensual" model that contains core features common to many of the specific implementations.

2.1. *The Base Model*

Computational models of visual attention share a very similar overall organization, which follows at least coarsely the structure and organization of the visual percep-

representation of target location   representation of target content

spatial map

feature maps

visual input

Figure 15.1. Base model depicting core elements of the visual system and their connections, within which attention is implemented.

tual system. This commonality enables the description of a core consensual model architecture, consisting of a set of primary cognitive elements that are present in different forms throughout a variety of computational models. The base model provides a joint reference for reviewing and comparing specific computational solutions proposed by the individual models. It consists of topographically and hierarchically organized feature maps, a spatial map coding locations, and two modules providing top-down bias by coding target location and content, respectively (Figure 15.1).

Feature maps are postulated to code for specific properties of visual input. Such feature maps were first proposed in the influential feature integration theory of attention (Treisman, 1999; Treisman & Gelade, 1980), which assumed that basic visual features are represented in separate topographical maps, each of them linked to a master map of locations that allows later binding of individual features for further processing. Feature maps, at least in their initial computational implementation, were assumed to code low-level basic features of visual input,

such as color, orientation, and intensity (Koch & Ullman, 1985) that could be represented in primary visual cortex (VI). Yet, it has also been appreciated that such maps might as well be used to code progressively more complex features, such as motion, shape, and object identity, depending on the focus and complexity of the model. In cognitive neuroscience terms, these maps represent different stages of visual analysis of the ventral, or "what," stream, starting from the primary visual cortex to the inferotemporal cortex (IT). Although the initial feature maps coding low-level features are supposed to be independent, their combination to ever higher levels of representation can be seen as comprising a hierarchical system dedicated to object recognition (Mozer et al., 1998). The maps are topographically organized with the nodes in the initial maps having relatively small reception fields, and the nodes in the higher feature maps having large receptive fields, ultimately covering the whole visual field.

Whereas feature maps deal with what is present in visual input, spatial maps code information about where visual input is present. Such coding occurs through topographical representation of locations in the visual field. In feed-forward models of attention, the spatial map also frequently serves as the map that explicitly codes spatial attention, earning the name saliency map (Koch & Ullman, 1985), activation map (Wolfe, 1994), or simply attentional map (Mozer et al., 1998). The spatial map is supposed to be instantiated in the dorsal, or "where," stream of visual processing, most frequently in the posterior parietal (PP) cortex. Feature maps and the spatial map are densely interconnected. Most models assume each of the feature maps to be connected to the location map. In many models, these connections are unidirectional, leading primarily from low-level feature maps to the spatial map (e.g. Koch & Ullman, 1985); however, other models include recurrent feedback connections, from the spatial map back to low-level feature maps (e.g., Deco, 2001).
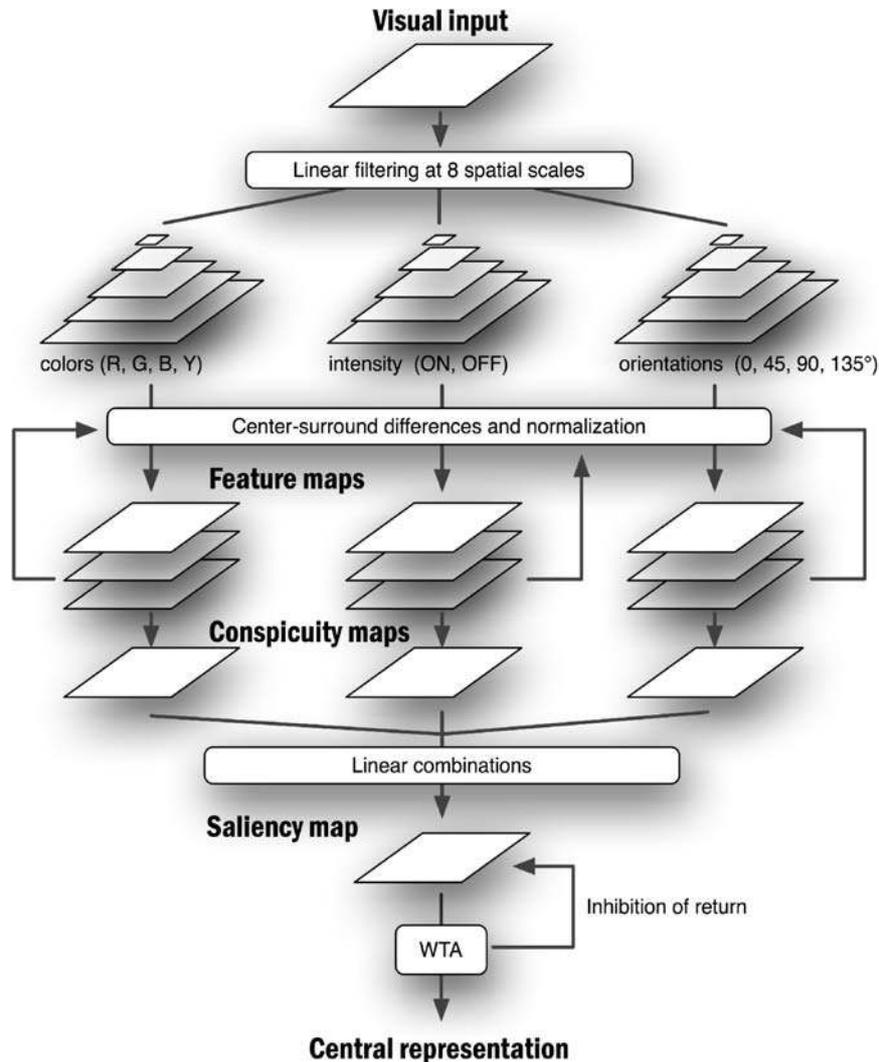
The feature maps and spatial map represent the most frequent core of the model, where attentional influences emerge and are expressed. To also model the top-down influence on attention, most models assume modulatory connections from structures coding goals and intentions. In the base model, the top-down modulatory effect is exerted by an element that is holding a representation of target location and an element that is holding a representation of target content. The former projects to the spatial map and the latter to the feature maps. Both are assumed to be located in the anterior part of the brain (i.e., in the prefrontal cortex [PFC]), closely connected to cognitive control processes and the production of goal-directed behavior.

2.2. Explicit Computation and Representation of Attention

Within the architecture of the base model, various computational models of attention can be implemented, differing significantly in the pattern of connectivity, the functional roles played by individual components, and the ensuing dynamics and behavior of the model. One conceptualization of attention assumes it is a distinct, explicitly computed and represented feature of the system enabling selection and filtering of visual input for further analysis. Attentional dynamics are assumed to evolve through two clearly defined steps. In the first step, stimulus features are used to compute and identify most salient locations in the visual field, representing the focus of attention. In the second step, the representation of spatial attention is used to focus the flow of visual information in an object processing stream.

Computation of visual attention using the saliency map was first explicitly proposed in the model of Koch and Ullman (1985) and led to a number of similar implementations in other models (Mozer et al., 1998; Wolfe, 1994) as well as more detailed refinements and additions to the original proposed mechanisms of feature extraction and saliency computation (Itti & Baldi, 2005; Itti, Koch, & Niebur, 1998; Lee et al., 1999). As proposed by Koch and Ullman (1985; see Figure 15.2), the visual input is first

Visual input

Linear filtering at 8 spatial scales

colors (R, G, B, Y)   intensity (ON, OFF)   orientations (0, 45, 90, 135°)

Center-surround differences and normalization

Feature maps

Conspicuity maps

Linear combinations

Saliency map

Inhibition of return

WTA

Central representation

Figure 15.2. A schematic representation of the saliency map based on the computational model of visual attention initially proposed by Koch and Ullman (1985) and fully implemented by Itti et al. (1998).

decomposed by several feature detection mechanisms working in parallel at different spatial scales. Resulting individual feature maps represent the salience of locations in regard to the basic visual features being represented. The key element in determining the saliency of a location is not the intensity of the feature, but rather its local contrast judged in the context of the rest of the visual field. A red dot is more likely to attract attention than a uniform field of red. Furthermore, a particular red dot is more likely to attract attention when

it is the only red dot in the visual field than when it is just one of many in a field of red dots. To take both properties of attention into account, the computation of saliency is proposed to be based on both short-range center-surround differences, which identify the presence of local contrast, as well as long-range spatial normalization, which estimates its importance in regard to the entire visual field. Resulting activity in feature maps is combined in "conspicuity maps" for each feature type and summed into a single, "master" saliency map coding overall

saliency of stimuli within a given location in a topographical representation of the visual field. In this manner, the model collapses the representation of saliency over specific visual features, making the map blind and indifferent to which feature caused specific locations to be salient.

Once saliency is computed, the model has to be able to select a single location on which to focus attention. This step in ensured through a separate winner-take-all (WTA) network. Receiving topographical input from the saliency map and implementing strong global inhibition, the WTA network quickly settles on the winning neuron (or a population of them) receiving highest activation from the saliency map and representing the focus of attention.

Given a static display, the described network would compute and lock on to the most salient location in the visual field. To be able to disengage from the winning location and explore other salient locations, the model has to incorporate an "inhibition of return" (IOR) mechanism that temporarily inhibits the activity in the winning location. This inhibition then enables the second most active location in the saliency map to drive the shift in activity in the WTA network, representing a new focus of attention. Such IOR in covert shifts of attention has been experimentally well demonstrated (Kwak & Egeth, 1992; Posner, Cohen, & Rafal, 1982). In the Itti and Koch (2000) implementation of the model, it is realized through inhibitory feedback from the WTA network back to the saliency map. Other models have been developed that use similar types of active or passive (e.g., fatigue-like) inhibition mechanisms (Houghton & Tipper, 1996; O'Reilly & Munakata, 2000). Depending on the parameters of the model, IOR enables a network to sequentially select or search through a number of the most salient locations in the visual field before returning to the initial one.

A number of implementations of the saliency-based computational models have shown it to be successful in predicting human performance in psychophysical experiments and visual search tasks (Itti & Koch, 2000; van de Laar, Heskes, & Gielen, 1997), as well as accounting for the pattern of human eye movements made during the viewing of images containing complex natural and artificial scenes (Parkhurst et al., 2002). For an excellent review of saliency-based computation of attention, see Itti and Koch (2001).

Once attention is focused on a specific location, the mechanism of guiding further visual processing needs to be specified. The most straightforward solution uses saliency representation as a gating signal modulating the flow of information from lower-level feature maps to higher levels of visual analysis. In a model proposed by Mozer et al. (1998), the information coming from low-level feature maps is multiplied by the activity in the topographically equivalent area of the saliency map, limiting further processing to salient locations while attenuating the rest. Significantly more complex solutions are based on a dynamical routing approach proposed by Olshausen, Anderson, and Van Essen (1993) and recently instantiated by Heinke and Humphreys (2003) in their Selective Attention for Identification Model (SAIM). The routing and SAIM models build on the idea that translation-invariant pattern recognition can be achieved by an attentional window that can be moved over the visual field, focusing on its relevant sections and feeding that partial image to a recognition network. The task is realized by a complex network of connections (termed the "contents network") that map retinal input through a number of stages to a smaller "focus of attention" (FOA) layer. The appropriate mapping is ensured by a selection network (spatial map analogue), whose mutually inhibitory units activate only those connections of the contents network that project from the currently relevant part of the input layer to the FOA layer. In this manner, the network not only gates the visual input, but also translates it to a single layer for further analysis. Both models are successful in accounting for a number of empirical findings relating to both normal as well as pathological attentional phenomena.

For a review, see Heinke and Humphreys (2005).

Although bottom-up influences are important in drawing our attention to objects in the environment, models of visual attention must also appropriately account for voluntary top-down control. The models considered so far allow for both spatially based as well as feature-based top-down control of attention. Intentional guidance of spatial attention is presumed to occur via top-down inputs to the spatial map, which either bias or directly determine its pattern of activation. On the other hand, feature- or object-based attention is assumed to be brought about by biasing the computation of saliency. Searching for red horizontal bars in the visual scene would, for instance, entail selectively enhancing the contribution of feature maps coding red color and horizontal orientation to the master map of saliency, leading to the highest buildup of activity in location(s) where a conjunction of both features is present (Wolfe, 1994). As recently shown by Navalpakkam and Itti (2005), top-down control of attention using biased computation of the saliency map is not limited to simple features. Using learned sets of low-level features related to different views of an object, their model was successful in locating complex visual objects in natural scenes.

2.3. *Interactive Emergence of Attention*

An alternative approach to conceptualizing attention has been to consider it an emergent property of the system evolving seamlessly through competitive interactions between modules. Dense bidirectional (i.e., recurrent) connections between processing modules enable the active representation in any module to be the source or the target of a biasing signal affecting the local competition between representations, hence, the name "biased competition models." This bidirectional connectivity enables a dynamic settling process to occur that stabilizes on a coherent representation expressed throughout the system. Attention is not computed explicitly through distinct steps but rather emerges continuously as a property of activation dynamics in the system (Desimone

& Duncan, 1995). The initial proposal for these types of models can be traced to Phaf et al. (1990) and Desimone and Duncan (1995), with more recent models being proposed by Ward (1999) and Deco (2001; Deco & Rolls, 2005a).
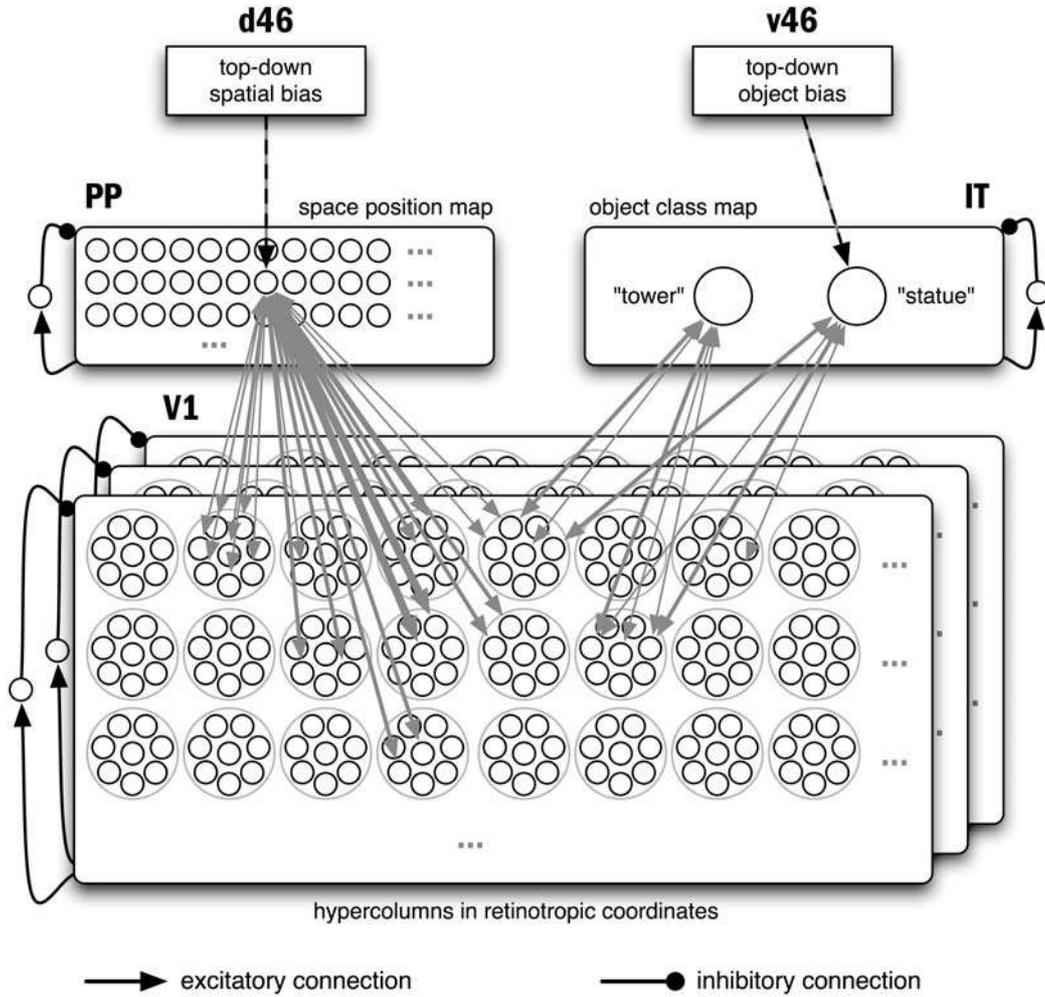
Representative biased competition models of visual attention have been developed, described and explored by Deco and colleagues (Deco, 2001; Deco & Lee, 2004; Deco, Polatos, & Zihl, 2002; Deco & Rolls, 2002, 2003, 2004, 2005a; Deco & Zihl, 2004). The simplest instantiation of the model assumes existence of three processing modules, V1, PP, and IT, respectively corresponding to low-level feature maps, the spatial map, and the high-level feature map in the base model (Figure 15.3). Each module consists of a number of units, each representing a pool of neurons with similar properties. The activity of each unit is described using mean field approximation, where each unit i is characterized by its activation x_i, reflecting an average firing rate of the pool and an activity level of the input current A_i. The input-output relationship is defined as:

$$x_i = F(A_i(t)) = \frac{1}{T_r - \tau \log(1 - 1/\tau A_i(t))}$$

in which T_r denotes the cell's absolute refractory period (e.g., 1 ms) and τ stands for the membrane time constant. The dynamics of each excitatory unit within a module is described by:

$$\tau \frac{\delta}{\delta t} A_i(t) = -A_i + aF(A_i(t)) - bF(A^I(t)) + I_i^B(t) + I_i^T(t) + I_0 + v.$$

The first term is a habituation decay term. The second term represents the recurrent self-excitation that maintains the activity of the cells and mediates their cooperative interaction within the unit ($a = 0.95$). The third term represents a local inhibitory input from the inhibitory unit providing the basis for local competition between excitatory units within the module ($b = 0.8$). I_i^B denotes a specific bottom-up input from a lower cortical module, whereas I_i^T represents a specific top-down bias from higher

Figure 15.3. Schematic representation of biased competition model (Deco, 2001). Units within V1 hypercolumns coding eight orientations at three spatial scales are connected to (PP) units bidirectionally with Gaussian spatial distribution of weights. Units in IT are bidirectionally connected to every unit in V1 with variable weights defined through supervised Hebbian learning rule. (Only sample connections with various weights are shown.)

cortical modules. I_0 denotes a spontaneous background input, and v is an additive Gaussian noise in the system.

The inhibitory unit integrates information from all the excitatory units within the module and feeds back nonspecific inhibition to all excitatory units in the module. Its dynamics is defined by:

$$\tau_I \frac{\delta}{\delta t} A^I(t) = -A^I - cF(A^I(t))$$

$$+ d \sum_{i=1}^{m} F(A_i(t)).$$

With $\tau_I = 7\,\mathrm{ms}$, $c = 0.1$, and $d = 0.1$, the first two terms describe decay and self-excitation, respectively, whereas the third term is a function of activities of all excitatory units within the module connected to the inhibitory unit.

Input I_i^S of any connected module that provides either bottom-up input or top-down bias is described by:

$$I_i^S(t) = \alpha \sum_{j=1}^{n} w_{ij} F\left(A_j^S(t)\right),$$

where A_j^S denotes the activity level of the source module unit, w_{ij} denotes the connection weight between source unit j and target unit i, and $1/\alpha$ reflects an attenuation factor. Setting $\alpha = 1$ for bottom-up input and $\alpha = 0.6$ for top-down bias prevents the latter from dominating V1 units, and allows lower level representations to change the state of higher-order modules. The V1 module represents the input layer of the visual system and consists of a lattice of 33×33 hypercolumns topographically covering a 66- $\times$ -66-pixel scene. Each hypercolumn consists of twenty-four excitatory feature detector units (pools) representing eight spatial orientations in three spatial scales. The sensory input to the excitatory units is computed using 2D-Gabor functions, which act as local spatial bandpass filters detecting the presence of sensory input in a given orientation at a given spatial scale in the relevant location of the presented input image (for additional details, see Deco, 2001 and Lee, 1996). An additional inhibitory unit per scale is used to mediate global normalization within units at each scale.

The PP module encodes spatial location in the visual field, representing the function of the PPC. It consists of a lattice of 66×66 units, each of them receiving input from a limited spatial neighborhood of V1 hypercolumns. To capture the Gaussian-like nature of the spread of activation, the mutual connection weights between the units at hypercolumn pq in V1 and unit ij in PP are given by equation:

$$w_{pqij} = Ce^{-\frac{(i-p)^2+(j-q)^2}{2\sigma_w^2}} - B.$$

With $C = 1.5$, $B = 0.5$, and $\sigma_w = 2$, the resulting center-excitatory, surround-inhibitory weight profile connects bilaterally each PP unit to a spatial neighborhood of about 5×5 V1 hypercolumns, giving an effective receptive field of about 17 pixels in diameter. Local competition between PP units is ensured by one inhibitory unit that receives input from all excitatory units and inhibits all units uniformly, enabling WTA competition within PP.

The IT module encodes object class or categorical information corresponding to the function of inferotemporal cortex. IT consists of a finite set of units, each receiving connections from all units in V1 and returning attenuated symmetrical reciprocal feedback connections to V1. Similarly to PP, local competition is ensured by an additional inhibitory unit receiving excitatory input from and returning inhibitory feedback to all excitatory units in IT. Connection weights between V1 an IT are trained by supervised Hebbian learning. During learning, a target image is presented as input to V1 whereas top-down bias is imposed on PP unit coding for location of the target and IT unit coding for its identity. The network is allowed to settle into a steady state, after which all the relevant V1–IT connection weights are updated using Hebbian learning rule:

$$w_{ij} = w_{ij} + \eta F\left(A_i^{V1}(t)\right) F\left(A_j^{IT}(t)\right)$$

where η denotes the learning coefficient and t is large enough to allow for convergence. Having successfully learned to perform translation invariant object recognition, the model can operate in three modes: preattentive mode, spatial attention mode, and object attention mode (Deco, 2001; Deco & Lee, 2004). In the preattentive mode, no top-down biasing signal is provided. The perceptually most salient object in the visual field will cause a stronger input from feature maps to the representation of its location in the spatial map. This enhanced activation of the relevant location in the spatial map is then fed back to the feature maps, biasing their activation and thus the flow of information to the object recognition pathway. The recurrent, bidirectional flow of activity occurring in biased competition models results in a positive feedback loop that leads to iterative convergence on a single winning representation both in the spatial map as well as in the feature maps and the object identity module. This winning representation effectively marks both the position and the identity of the most salient object in the visual field.

In the spatial attention mode, object recognition at the attended location is

implemented through preselecting a particular location in the representation of target location, presumably hosted in the dorsal PFC, which provides a top-down bias causing the activation of appropriate units in the spatial map. Feedback connections from the spatial map to the low-level feature maps enhance the activation of corresponding units, acting as a spatial attention beam. Excited by both the sensory input and the top-down signal from the spatial map, those units representing the features at the attended location will provide a stronger input to the related identity units in the higher-level maps, enabling them to win the local competition leading to identification of the object present at the attended location. In a model simulation, implementing a top-down spatial bias to PP results in an early differentiation of activity between the cell assemblies coding the target and distractor locations, respectively. The differentiation spreads both to V1 and IT, causing the cell assemblies coding the target object to be significantly more active than the ones coding the distractor object, signaling object identification.

The dynamics of the "object attention" mode mirrors that of the "spatial attention" mode. A biasing signal arising presumably from the ventral PFC leads to activation of the higher-level map units coding the identity of the attended object. Feedback connections to the low-level feature maps enhance the activity of units coding visual features of the attended object, effectively "back-projecting" the response pattern associated with the object across all retinotopic locations in parallel. The units receiving input from an appropriate visual stimulus will resonate best with the feedback signals leading to their enhanced activation. Providing stronger input to the units in the spatial map that code the position of the attended object will enable them to win the local competition, effectively completing visual search. Monitoring the dynamics of the model simulation reveals that the local competition is first resolved in the IT module, which then drives the competition in V1 and PP modules in favor of the units corresponding to the target object. The object is considered

found when competition is eventually resolved in PP.

Simulations using the biased competition model were found to be successful in accounting for a number of empirical results in visual search (Deco & Lee, 2004; Deco & Zihl, 2001b). For example, the model showed that added difficulty of constraints in conjunction search tasks causes the network to take longer to settle. Congruent with behavioral findings, the times for the network to settle were independent of the number of distractors in a feature set task, whereas the times in conjunction search tasks were progressively longer with increasing number of distractors. Furthermore, reaction time slopes related to different types of conjunction search obtained by model simulations were successful in predicting subsequent psychophysical investigations (Deco et al., 2002). The model thus demonstrated that some, seemingly serial cognitive tasks may actually be a result of neuronal dynamics in a fully parallel system, bypassing the need for a dedicated implementation of a serial process guiding attentional spotlight from one item to the other.

Introducing artificial lesions in the model, enables testing of possible accounts of attentional deficits caused by brain lesions. Selective damage to the right side of the PP module reproduced some of the symptoms of the left spatial hemineglect typically caused by lesions to the right parietal cortex (Heinke et al., 2002). Replacing global inhibition with the local lateral inhibition enabled the model to also account for object-based neglect in which only the left side of the objects in the visual field is not seen (Deco & Rolls, 2002). Additionally, it also provided novel predictions about how patients with object-based neglect might perceive objects when they are joined with cross-links or brought toward each other (Deco & Rolls, 2002).

2.4. Key Issues in Models of Visual Attention

The present overview of visual attention models offers only selected highlights of some of the important progress being made in recent years. Advances in understanding
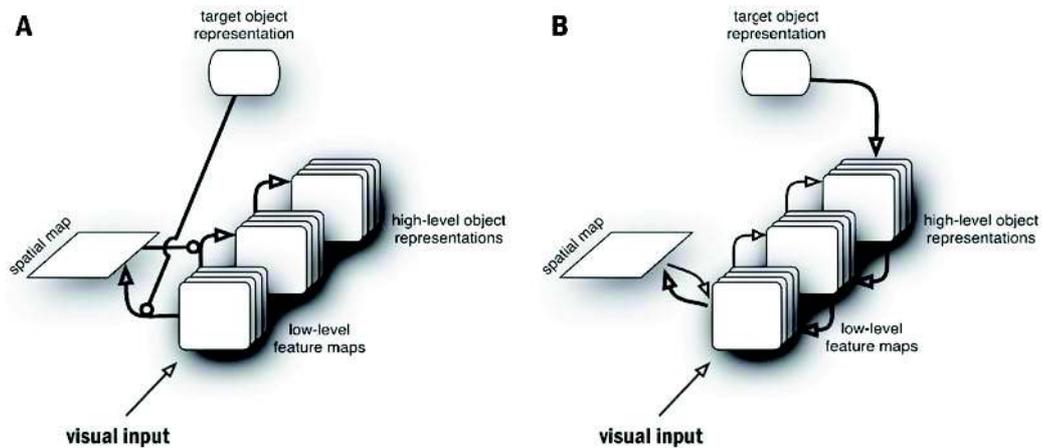
Figure 15.4. Two possible routes of top-down object attention. (A) In feedforward models, attention is guided by biasing input from the low-level feature maps to the spatial map that gates the flow of information to the higher levels of visual processing. (B) In biased competition account, the attention is guided by propagation of bias over rich recurrent connections from a high-level object representation to the low-level feature maps and through them, to the spatial map, ultimately leading the network to settle in a state representing the target object and its location.

of the architectural structure of the visual system have enabled the design of computational models that closely mimic the known neurophysiology of vision and are able to qualitatively match a wide variety of neurophysiological findings. They also agree with behavioral results coming from the basic experimental paradigms and with the data from brain-damaged patients suffering from attentional impairments. Nevertheless, despite convergence in a number of areas, important dilemmas still remain, most of them illustrated by the differences between the two basic categories of attentional architecture described earlier: feedforward versus biased competition (see Figure 15.4).

The core dilemma relates to the question of representation of attention: Should attention be represented explicitly in a single or perhaps in multiple spatial maps that code saliency of visual areas, or should it be represented implicitly in the interactive dynamics of the network? The first alternative is embodied in the models of attention centered around a feed-forward saliency map mechanism, which can be traced back to the proposal by Koch and Ullman (1985). In focusing on the problem of effective computation of saliency, these models have been effective in capturing the known neurobiology of

low-level visual processing, while simulating findings from the empirical visual search and natural scene viewing, and providing a successful architecture for various computer vision applications.

The second alternative builds on the conceptualization of attention as an emergent property of activation dynamics. It relies on rich recurrent connections between processing modules that bias local competition between representations. Although successful in replicating visual search findings, the true strength of these models lies in their ability to model the qualitative pattern of impairments associated with neuropsychologically based attentional disorders, such as the spatial neglect syndrome (Deco & Rolls, 2002), and in providing a coherent and seamless neural architecture that relates perception to action (Ward, 1999).

Both types of models exhibit a range of specific strengths and weaknesses. Their future development will depend on their ability to relate to the known brain anatomy and physiology (Shipp, 2004). In this regard, those models that incorporate a detailed mathematical description of neuronal dynamics are already successful in replicating and predicting spiking activity of single neurons (Deco & Rolls, 2003, 2005b), as

well as local population dynamics, as reflected in the hemodynamical response observed with functional magnetic resonance imaging (fMRI; Deco et al., 2004). Besides the models presented here, there are also other contenders providing alternative approaches that should be considered, among them, Bundesen's Theory of Visual Attention (TVA). It started as a formal mathematical theory describing behavioral results (Bundesen, 1990, 1998), but was recently developed into a neural theory (Bundesen, Habekost, & Kyllingsbaek, 2005), which successfully applies the same basic equations to provide both a quantitative account of human performance on a set of attentional experimental paradigms and an account of a range of attentional phenomena studied at the single cell level using electrophysiology.

Furthermore, the models of visual attention will also need to successfully scale up toward the more complex visual tasks, including higher-level cognitive processing (Deco & Rolls, 2005a; Navalpakkam & Itti, 2005). Moreover, to provide a comprehensive description of visual attention and to eliminate any remnant of the "ghost in the machine," the models of attention in visual processing will have to be related to those explaining the ways in which top-down, goal-driven intentions are represented, manipulated, and controlled. These models have been developed and explored within the research of cognitive control, which will be addressed in the following section.

3. Models of Goal-Driven Attentional Control

For many theorists, the terms executive control, cognitive control, controlled attention, and executive attention are interchangeable (and they will be used somewhat interchangeably here as well), referring to the notion that sometimes attention appears to be directed in a top-down, volitional fashion according to abstract, internally represented goals, rather than by detection or extraction of specific perceptual features or objects. Similarly, in some cases, attention appears to have its effect in biasing the selection of

actions rather than inputs, or more globally, in modulating whole task-processing pathways rather than specific components of perception.

Cognitive control is often described in opposition to automaticity. Automaticity refers to the capacity of a cognitive system to streamline well-practiced behavior, so that task-relevant actions can be executed with minimal effort. As a complement to automatic behavior, cognitive control refers instead to the effortful biasing or inhibiting of sensory-motor information in the service of novel and unpracticed goal-directed behaviors. Top-down attention is what arises out of the neuronal activity shift guided by cognitive control, and it is typically assumed to be the product of biasing representations (such as intentions, rules, goals, and task demands) in the PFC that compete with perceptually based representations in the posterior cortex. Cognitive control is the mechanism that guides the entire cognitive system and orchestrates thinking and acting, and top-down attention is interpreted as its main emergent consequence.

Computational models seem best positioned to describe how top-down attentional control is engaged during the course of task processing and to indicate the consequences of such engagement. Critically, the explanations that arrive out of computational models are explicitly mechanistic in character, and they minimize the reliance on a hidden homunculus. Although formal theoretical investigations in the study of cognitive control have not advanced to the same degree as those in visual selective attention, there have been a number of computational models developed in this domain. Many of these models adopt the biased competition framework discussed in the preceding section as a core architectural assumption. Additionally, a primary focus of most models has been to address human experimental data, arising from basic cognitive performance, neuropsychological impairment, and neuroimaging findings, particularly regarding PFC function. This may be because many of the core phenomena of cognitive control relate to tasks most easily examined in humans, although this
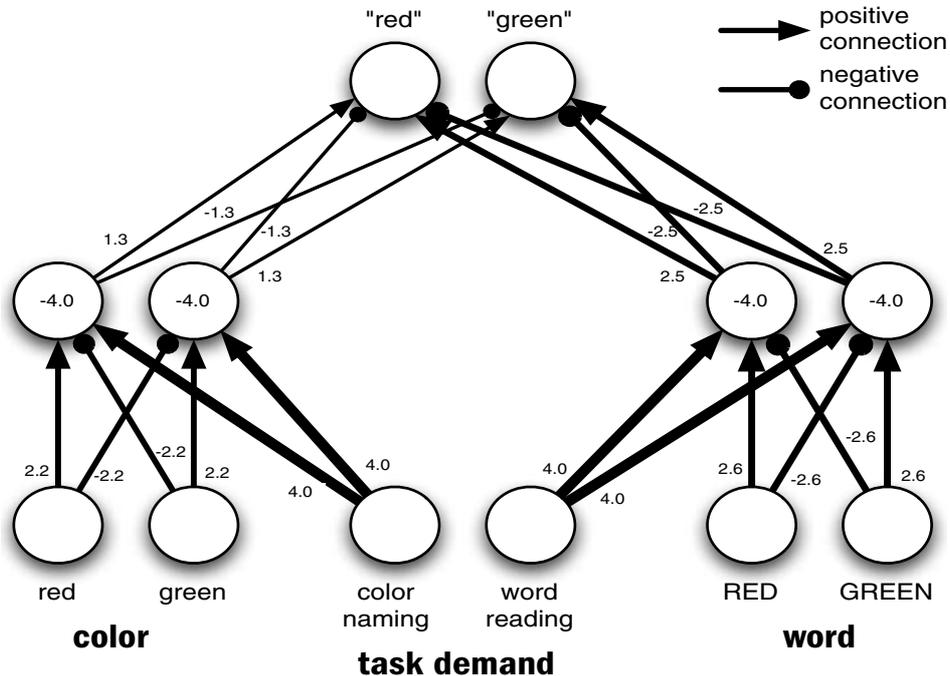
Figure 15.5. J. Cohen's model of the Stroop test (Cohen et al., 1990) This model provides a minimal account of top-down attentional biasing effects emerging from prefrontal cortex-based task-set representations.

has begun to change more recently (e.g., Miller & Cohen, 2001). In the next section, a fundamental but minimal model that illustrates the core principles of cognitive control is described first, then other key cognitive control models and the attentional issues they address are discussed.

3.1. *The Base Model*

A basic model that illustrates the key hypothesized mechanisms of attentional control is one developed by Cohen, Dunbar, and McClelland (1990) to account for processing and behavioral performance during the classic Stroop test (Stroop, 1935) of selective attention. The Stroop test may represent the paradigmatic example of the relationship and contrast between automaticity and cognitive control. The basic paradigm (although there have been many different variants) involves processing of colored word stimuli and selectively attending to either the word name or ink color. Attention is thought to be more critical for color nam-

ing than word reading, because the latter skill is so highly overlearned and practiced for most literate adults. The role of attention is especially critical for color naming in incongruent trials in which there is a direct conflict between the ink color and the color indicated by the word name (e.g., the word "GREEN" in red ink). In such a case, cognitive control over attention must enable preferential processing in a weaker task pathway (color naming) over a competing and stronger but task-irrelevant one (word reading).

The Cohen et al. (1990) model put forth a highly influential framework for understanding the mechanisms of cognitive control and attention in the Stroop task. Critically, the model illustrates very simple principles of biased competition in that attention is just another source of input that serves to strengthen the activation of hidden layer units, which then leads to a shift in the outcome of competition within a response layer (see Figure 15.5). The original model is feed-forward, although later models have

used a fully bidirectional architecture (Cohen & Huston, 1994; O'Reilly & Munakata, 2000) that includes more natural lateral inhibitory mechanisms. The model uses a standard connectionist activation framework in which the activation a_j of each unit j at time t is a logistic function of the net input:

$$a_j(t) = \frac{1}{1 + e^{-net_j(t)}}.$$

The net input from every unit i into unit j is first computed as:

$$rawnet_j(t) = \sum_i a_i(t)w_{ij}$$

where w_{ij} is the weight from each unit i to unit j. This raw net input is then transformed into a "cascade" form (McClelland, 1979) to simulate continuous time dynamics:

$$net_j(t) = ((1 - \tau)^*net_j(t - 1))$$
$$+ (\tau^*rawnet_j(t))$$

where τ is a constant.

Attention demands arise in the model because of the asymmetry of weight strengths in the word-reading versus color-naming task pathways. This asymmetry arises during a learning phase in which the network receives greater practice in word reading than color naming. Because the training phase is accomplished with the backpropagation learning algorithm (Rumelhart & McClelland, 1986), weight strengths change in direct proportion to training experience. The key attentional mechanism arises from the task demand units, which represent top-down attentional effects arising out of the PFC. These units have a sensitizing effect on hidden-layer activation, particularly for the color pathway, such that with task demand (attentional) input, the color hidden units are maximally sensitive to stimulus input and can compete strongly with activation arising out of the word pathway. The magnitude of the attentional effects depend on the size of the weights from the task demand units to the hidden layer, and they are computed as a cascading net input defined as in the previous equation.

Another core principle behind this model, which was also present in the original Norman and Shallice (1986) theory of cognitive control, is that the attentional system does not directly enable task processing, but only modulates its efficacy. This can be illustrated in the model in that the two task pathways, representing word reading and color naming, can each work in isolation (i.e., for unidimensional stimuli) to produce task-appropriate processing and responses, even in the absence of attentional signals. However, when both the word-reading and color-naming pathways are simultaneously engaged, competition between the two dimensions that occurs at the level of overlapping response representations produces a demand for attentional intervention. This demand for attention is most acute when performing color naming under competition conditions, because of the weaker strength of the color pathway. Thus, in the absence of attentional modulation, the word-reading pathway will dominate processing competition at the response layer.

In contrast, when there is an attentional influence from the task demand input on the color pathway, this pathway can successfully compete with the otherwise stronger word pathway by providing a stronger input to the response layer from the color-naming hidden layer. In the model, the mechanism of attentional modulation occurs via a nonlinearity of the activation function in task-processing units, such that, under the influence of top-down control, the activation function will be in its most sensitive region to be activated by bottom-up input, whereas without such an influence, the sensitivity to input is greatly reduced. Such top-down biasing mechanisms cause the color pathway to be more sensitive to the presence of color stimuli. This effect leads to a shift in the outcome of competition such that the color dimension successfully drives the response.

It is important to note that in the Stroop model, attention serves as an emergent

influence in that the activity in the task demand unit has a top-down biasing effect on the information processing taking place in the rest of the network. But this top-down biasing role does not have any special property, that is, these higher-level units are conceptually identical to the other units in the network. Therefore, attention is framed as a very general property that can arise out of the influence that representations of any kind can have on processing of information taking place in any other area.

A further postulate of the Cohen et al. model of attentional control, which was further elaborated in later papers (Cohen, Braver, & O'Reilly, 1996; Cohen & Huston, 1994; Cohen & Servan-Schreiber, 1992), goal-driven attentional biasing effects are critically related to the functions of PFC. In this region, goal-related contextual information is thought to be actively represented and feeds back into other regions of the posterior neocortex, where it can exert a top-down bias on competitive interactions occurring among local populations (Miller & Cohen, 2001). As a consequence of this coordinative activity, the PFC can both implement a top-down sustained attentional function (to keep active and operate on representations elsewhere in the brain) and also an inhibitory one (to suppress task-irrelevant pathways), but with this latter function emerging as an indirect consequence of excitatory attentional bias on local competitions, rather than via a direct top-down inhibitory signal.

3.2. Extensions and Alternatives to the Base Model

This basic mechanism of PFC-mediated top-down attentional biasing that forms the core of the Stroop model has provided a relatively comprehensive and influential account of a range of empirical phenomena. Moreover, the same architectural framework has been utilized to simulate a range of other attentional phenomena in the Stroop task and in other attention and cognitive control paradigms (e.g., Barch et al., 1999; Braver

& Cohen, 2001; Carter et al., 1998; Cohen et al., 1994; Dehaene & Changeux, 1991; Servan-Schreiber et al., 1998). A recent extension of the basic model was utilized to address fMRI data regarding the activation of PFC and posterior cortical regions during Stroop performance (Herd, Banich, & O'Reilly, 2006). A key feature of this recent model was the addition of a separate task demand unit coding for general color-related representations, both perceptual and linguistic. In other studies using the Stroop model as a theoretical framework, the primary motivation was to investigate the cognitive impairments in schizophrenia, a psychiatric condition believed to involve impairments of cognitive control due to alterations in the transmission of dopamine in the PFC. Individuals with schizophrenia, for example, are well known to show particularly large interference effects in the Stroop task, although recent data have suggested that the empirical phenomena are more complex than originally thought (Barch, Carter, & Cohen, 2004). The Cohen et al. (1990) model suggests that weakened attentional representations in schizophrenia patients impair the ability to successfully bias competition in favor of color naming over word reading, even when required by task conditions (Cohen & Servan-Schreiber, 1992).

Given the role of the Stroop task as the paradigmatic example of selective attention, it is perhaps not surprising that a variety of alternative computational models have been developed to explain attention in the Stroop. Yet, in many ways, these alternative models, which have been developed in both connectionist and symbolic architectures, can be seen as being formally very analogous in terms of attentional mechanisms to the Cohen et al. (1990) account. However, some of the models have had different emphasis, such as to try to explain Stroop phenomena within more generic and comprehensive architectural frameworks, such as modeling of visual attention more broadly (Phaf et al., 1990) and word reading (Roelofs, 2000), or to

account for potential high-level strategic variability (Lovett, 2002).

However, another recent model, put forth by Melara and Algom (2003), may provide an important conceptual alternative to the Cohen et al. (1990) Stroop model. In this so-called tectonic model, Stroop attentional effects are conceived of as being due to a continuous process of experience-dependent learning within two memory-based structures (the name *tectonic* for this theory, from the ancient Greek word *tektonikon*, meaning *to structure*, is due to this central feature of the model). One structure is a short-term memory of the dimensional uncertainty of the most recent trials, where values along the word dimension are more varied perceptually than values along the color dimension. The other structure is a long-term memory of the dimensional unbalance, storing asymmetry in the record of the observer's past efficiency in accessing the target dimension relative to the distractor dimension. This structure reflects the relative difficulty with which the currently accessed representations can be activated in long-term memory. Each structure contributes to building up excitation of the task-relevant dimension and inhibition of the task-irrelevant dimension. This complex model (see Melara & Algom, 2003, for equations and all technical details) has been shown to account for an impressive set of empirical behavioral phenomena that extend from the standard Stroop findings to other related effects, such as Garner interference. Nevertheless, the differences in the models may relate not to top-down attentional mechanisms per se, but to their interaction with a dynamically changing perceptual representation. Further work should be conducted to test the relationship between the Cohen et al. (1990) and tectonic model framework more systematically.

Subsequent models have attempted to expand the scope of the basic Stroop account by addressing the issue of the relationship of attention to the related construct of working memory. In particular, Cohen, Braver, and colleagues developed a model that integrated top-down biasing with the well-established active maintenance functions of PFC and also attempted to more thoroughly capture both the facilitation and inhibition effects of attention (Braver, Cohen, & Barch, 2002; Braver, Cohen, & Servan-Schreiber, 1995; Cohen et al., 1996). In this model, the central role of PFC is still to adapt the behavior of the entire cognitive system to the task demands via active representation of goal-related context, but additionally, the later models incorporated explicit mechanisms by which PFC representations could be actively maintained over time. Thus, in these models, top-down attentional effects could emerge following a delay interposed after presentation of a contextual cue. A further feature of this work was explicit incorporation of dopamine-mediated neuromodulation of PFC representations (Barch & Cohen, 1999; Braver & Cohen, 2000; Braver et al., 1995; Cohen, Braver & Brown, 2002). This dopamine modulatory input served both to stabilize active maintenance processes (via tonic dopamine activation in PFC) and to enable appropriate updating of PFC representations (via phasic dopamine activation, synchronous with cues indicating a new task goal or context). Other recent work has explored how norepinephrine neuromodulation, in addition to dopamine, might also play a particular role in modulating attentional focus (Aston-Jones & Cohen, 2005; Usher & Cohen, 1999; Yu & Dayan, 2005). More recently, a number of other models have been developed by distinct groups of investigators to address similar issues, but with more biologically detailed and realistic computational architectures (e.g., spiking units, distinct synaptic currents; Brunel & Wang, 2001; Durstewitz, Kelc, & Gunturkun, 1999). Nevertheless, these models have converged on similar prinicples regarding the role of biased competition mechanisms, active maintenance in PFC, and also dopaminergic neuromodulation in accounting for attentional effects in Stroop-like and other selective attention and working memory paradigms (for reviews of

this work, see Cohen et al., 2002; Durste-witz et al., 2000; O'Reilly, 2006).

Another key issue first addressed by the Stroop model but expanded in subsequent work is the role of inhibitory mechanisms in attention. In the Stroop model and many other biased competition based models of top-down attentional control, inhibition effects emerge as an indirect consequence of local competition, rather than as a direct explicit inhibitory mechanism. Yet, it is still controversial as to whether or not top-down attentional mechanisms might include a special inhibitory function, at least in some cases, such as in response inhibition tasks (Aron & Poldrack, 2006). There have been computational models developed that postulate a specialized role for direct attentional inhibition mechanisms as an alternative to the standard biased competition account as a means of explaining distractor suppression and negative priming type effects (Houghton & Tipper, 1996). However, even in this model, there is no "central" top-down inhibition mechanism; rather, the inhibitory effects are achieved by local positive and negative feedback circuits thought to be widely distributed throughout the brain.

A final area of recent activity in elaborating on the computational mechanisms of goal-driven attention concerns mechanisms by which attentional biases arise or are modulated during the course of task performance. In particular, one influential account has suggested that top-down attentional biases are modulated in response to mechanisms that monitor dimensions of ongoing performance. Specifically, it has been postulated that the anterior cingulate cortex (ACC) detects response conflict present during task performance and translates this conflict index into an output signal that modulates attentional biases within lateral PFC (Botvinick et al., 2001). The basic hypothesis is that when high conflict occurs between different motor or behavioral responses, cognitive control mechanisms intervene to bias the relevant response versus the others, thus overcoming the conflict. These interactions have been characterized

in terms of a single conflict-control loop mechanism, where the performance of certain task conditions leads to detection of response conflict, which in turn leads to the engagement or increase of cognitive control, and in improved conflict resolution in subsequent performance. However, a new model proposes that ACC–PFC interactions are described by two, rather than one, distinct conflict-control loops (De Pisapia & Braver, 2006). The first loop implements a reactive control mechanism in which conflict detected in ACC over a short-time scale transiently modulates PFC activity to adjust within-trial attentional biases. The second loop implements a proactive control mechanism, in which long-time scale conflict is also detected in ACC and more slowly adjusts attentional biases in PFC across trials. The model was used to successfully account for detailed aspects of behavioral and brain activation patterns in the Stroop task.

3.3. *Multi-Tasking*

The previous sections described computational models of attention that operate at different levels of information processing, from fine-grained influence on visual perception to representations of goal information actively maintained in working memory. An intriguing question that has recently been garnering a great deal of theoretical interest is whether there are even higher forms of attention, such as those that can aid in the selection of one out of many possible tasks to perform. In particular, the question is whether the attentional biasing effects discussed previously can operate not just at the level of perceptual features (e.g., red vs. green colors) or dimensions (color vs. word), but that can also influence the activation of whole task pathways over competing pathways. This issue becomes more clear when considering multitasking situations, which seem to approximate well the real-world demands of everyday cognition. In the so-called multi-tasking situations, more than one task needs to be performed at a time, either through simultaneous (i.e., nested or interleaved) engagement or through rapid
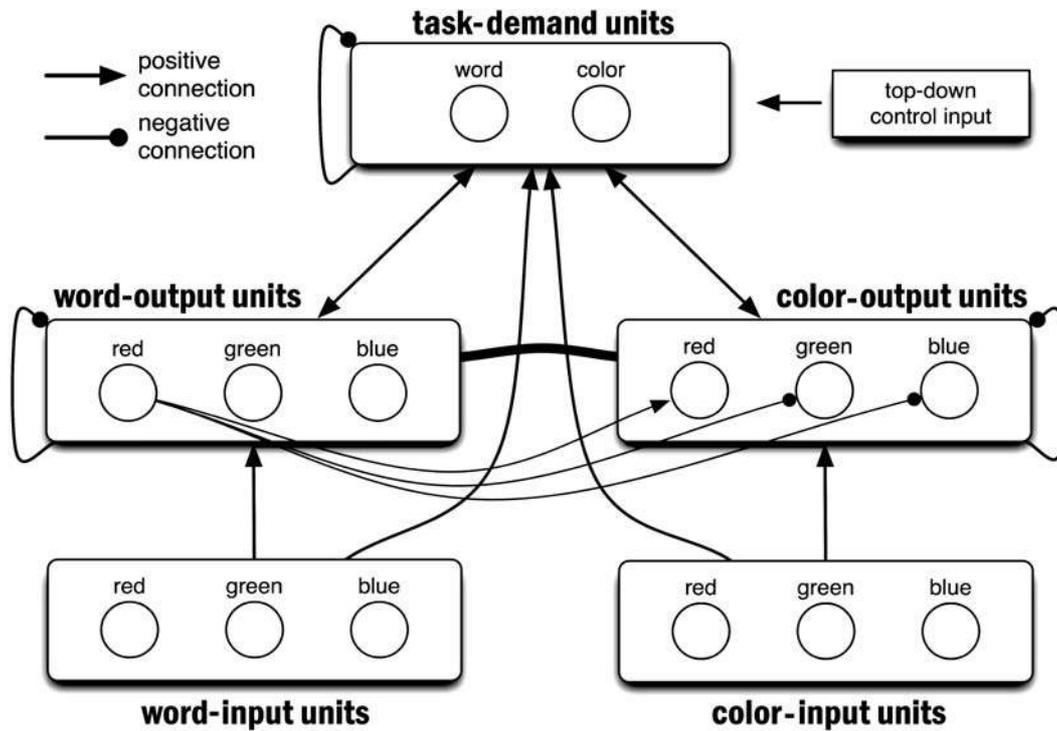
Figure 15.6. The Gilbert and Shallice (2002) model of task switching. This model is built on and extends earlier connectionist models of the Stroop task (Cohen et al., 1990; Cohen & Huston, 1994).

sequential alternation. Such situations seem to pose heavy attentional demands, and therefore they provide an excellent test bed for cognitive theories on attention at the task or dimensional level, rather than at the featural level of the stimuli.

One particular focus has been on *task switching*, an experimental paradigm requiring rapid switching among two or more tasks, in either an uncued-but-predictable or cued-but-random sequence. One consistent finding of such task-switching experiments is that there are reliable and robust *switch costs*; for example, performance is poorer on a trial in which the task is switched, in terms of both longer reaction times and a higher percentage of errors, compared with when the task is repeated. The original explanation for this finding is that a special time-consuming internal reconfiguration process is required to switch between tasks, which enables the engagement or "loading in" of the appropriate task representations that can bias attention appropriately during task per-

formance. However, an important question is whether this task-set reconfiguration process actually requires a dedicated mechanism that enables the appropriate form of attentional shift. A few theoretical models have been developed that provide an account of the types of attentional control and reconfiguration mechanisms involved in task switching.

One influential theoretical account of task switching that has been instantiated as a computational model is that of Gilbert & Shallice (2002; see Figure 15.6), which adopted the basic interactive architecture used in later models of the Stroop task (Cohen & Huston, 1994; O'Reilly & Munakata, 2000). The network consists of two separate input and output layers for words and colors, and a task-demand layer. In addition to top-down attentional effects, the task-demand units receive bottom-up connections from the input layers and the response layer. These bottom-up inputs allow for associative learning effects

and item-specific priming based on past experiences. The task-demand layer has one unit for the color-naming task and one unit for the word-reading task. Thus, the model has the potential to implement task-switching paradigms by shifting which task-set unit is active. Lateral inhibition between task pathways provides a means for top-down excitatory input to bias the outcome of representational competition. Task-demand units receive an input from a top-down control input, which specifies which of the two tasks the network has to execute for a particular trial.

All units in the model compute activations in response to the weighted sum of all incoming inputs, both top-down and bottom-up. Thus, as in the Cohen et al. (1990) Stroop model, there is no distinction between task-demand (attentional) input and bottom-up perceptual signals. The activations themselves are computed as in the standard interactive activation equations (McClelland & Rumelhart, 1981), where the increase in activation for each cycle is given by:

if the net input is positive:
$$\Delta act = step^*net^*(max - act)$$

if the net input is negative:
$$\Delta act = step^*net^*(act - min)$$

where *step* is the step size (establishing the speed of the activation update in each cycle), *net* is the net input, *max* is the maximum activation value allowed, and *min* is the minimum activation value. A random Gaussian noise is also added to the activation values of each unit.

The core feature of this model is that the task-demand units retain a residual level of activation even after that task is completed. This type of mechanism implements a version of the task-carryover account (Allport, Styles, & Hsieh, 1994), which postulates that switch costs are due to interference between this residual task-set activation and the engagement of a new task-set representation corresponding to the cur-

rently relevant task. Importantly, however, the model suggests that there is no specialized reconfiguration mechanism that is only engaged on switch trials. Activation of the relevant task-set representation occurs in the same way on every trial; it is just that on switch trials, there is increased competition between this new representation and the residual activation from the previously engaged task representation. Such competition does not occur on task-repeat trials when the same task-set representation as the previous trial is activated again.

A second important attentional mechanism implemented in the Gilbert and Shallice (2002) model is the bottom-up activation of task-set representations from features of task stimuli. The model implements a Hebbian (i.e., activity-dependent) learning mechanism. The weights between the stimulus input j and task demand units i are set with the learning rate *lrate* according following equation:

$$w_{ij} = lrate^*a_j^*a_i$$

This equation does not establish an update of the weights based on previous values, but instead the weights are calculated as new at the end of each trial, and the weights derived only affect the model's performance in the next trial. This mechanism allows the learning of associations between an active task-set representation and the particular stimulus features present on a task trial. This learning effect means that if such features are presented again on the subsequent trial, they will have the ability to "prime" the previously associated task-set representation due to strengthened associative weights.

Gilbert and Shallice (2002) applied their model to a task-switching version of the Stroop task (in which word reading and color naming randomly alternate across trials). The model was able to account for a wide range of experimental phenomena, including not only switch-cost effects and their temporal dynamics, but also phenomena related to item-specific priming effects as well. Importantly, the model's ability to

account for task-related attention does not rely on any type of specialized representations or mechanisms, but instead generalized mechanisms of biased competition, which play out not only in task-specific processing layers but also within the task-demand layer. Moreover, because the model is fully bidirectional and interactive, attention effects are fully emergent and can arise not only because of task-demand inputs, but also via effects emanating from the input level. However, one limitation of the Gilbert and Shallice (2002) model is that it does not address the question of whether task-demand representations themselves involve specialized content or coding schemes or how such "global" representations develop. This issue is taken up again in the concluding section of the chapter.

The Gilbert and Shallice (2002) model provides a useful starting point for understanding some of the core issues regarding computational mechanisms of task switching. In the last few years, other researchers have begun developing models that address some additional key issues in this literature. One issue concerns the mechanisms of task-set updating and advance task preparation and biasing. Some models have suggested that updating and advance preparation may occur in an all-or-none manner, but probabilistically across trials (Reynolds et al., 2006), or across the preparation interval (Sohn & Anderson, 2001). Interestingly, the Reynolds et al. (2006) model also makes contact with earlier models by postulating that dopamine neuromodulation is the source of the task-set updating and maintenance signal. In a second set of mathematical models, Logan and colleagues have argued that it is not necessary to postulate mechanisms of advance preparation and that task switching can occur purely retroactively as a retrieval process driven by the target presentation (Schneider & Logan, 2005). A different issue that has been addressed is the role of higher-order sequential processes in task switching (Brown, Reynolds, & Braver, 2006). In this work, conflict-control loops similar to those postulated in the previously described models of ACC–PFC interactions (i.e., Botvinick et al., 2001) adjust both attentional biases and response speed across trials in response to the experience of interference due to either task switches or the processing of task-irrelevant features.

3.4. *Dual-Task Coordination*

A second important component of multitasking arises in dual-task conditions, where two tasks must be performed in an overlapping period of time, such that some coordination or time sharing of processing resources is needed. Within this latter domain, there has been a great deal of interest in the so-called psychological refractory period (PRP) paradigm. In this dual-task paradigm, the relative timing of the two tasks is strictly controlled by manipulating the onset time of the target stimulus for the second task (termed T2) relative to the timing of the first task stimulus (termed T1). The basic finding is that that when the T2 onset time is short (relative to T1 reaction times), this causes an additional slowing of T2 reaction time (but not T1), which is termed the PRP effect. The primary theoretical interpretation of this effect is that there are certain stages of task processing that are particularly sensitive to dual-task overlap or interference and that some form of coordination mechanism is invoked to "serialize" processing as a means of minimizing this interference (Pashler, 1994).

Logan and Gordon (2001) developed a formal mathematical model that accounts for dual-task situations and used the model to extensively investigate the PRP effect. They found that only a model with a dedicated attention-switching mechanism was capable of fitting the behavioral data, thus suggesting that some form of reconfiguration of attention control takes place in overlapping task situations. Specifically, according to the model, a task-set refers to a particular set of control parameters that govern strategic aspects of how task stimuli are processed. When processing must rapidly shift from one task to another, new control

parameters have to be loaded, which may take a fixed amount of time. Ironically, as discussed earlier, this attention-switching mechanism was not found to be necessary to account for switch-cost effects in explicit task-switching paradigms (Logan, 2005). This is because the model does not assume any form of persistent storage of old control parameters after a task trial has been completed. In other words, according to the model, task-set switching processes will only affect performance latency when two tasks are overlapping in time.

Kieras and colleagues (Meyer & Kieras, 1997) have also modeled PRP effects in multitasking situations using the Executive-Process Interactive Control (EPIC) symbolic computational architecture. Their account of PRP phenomena assumes that dual-task coordination is purely under strategic control and that any form of serializing, time-sharing, or switching processes are not mandatory for performance. Instead, scheduling and task deferment is introduced in PRP situations to avoid potential interference or ordering confusions between tasks (e.g., responding to T2 before T1). Task deferment is accomplished by activation of a time-consuming control mechanism that implements lock and unlock commands on T2 processing. Thus, the engagement of this control mechanism is the source of PRP effects. Nevertheless, a key aspect of the model is that the point at which further T2 processing is "locked-out" (and then "unlocked" again) can depend on complex relationships between the two tasks and other experimental demands. Such flexible deferment implies an attentional control system that is strategic and subject to adjustments based on task experience. Typically, in simulations with the model, the primary determinant of changes in task-scheduling and deferment strategies is the presence of response level conflict or cross-talk. Thus, although as yet unexplored, the EPIC model may provide an account of the PRP effect that relies on dynamic conflict-control loop mechanisms similar to that postulated in the conflict monitoring account.

Other symbolic architectures, such as ACT-R, have also addressed the issue of dual-task coordination and PRP effects (Byrne & Anderson, 2001). In the ACT-R framework, the different modules – containing production rules – are intrinsically serial. Thus, in overlapping dual-task situations, ACT-R naturally implements a task-processing bottleneck that can induce PRP-type slowing. The inherent seriality of ACT-R sets it apart from both EPIC and other cognitive architectures in the treatment of dual-task attentional control.

A final issue that is just beginning to be explored is the more generic role of task-scheduling processes during multitask environments. The critical problem is that many multitask situations require a continuous and repeated inter-leaving of processing across different tasks due to the tasks' complexity and duration. Thus, in addition to the problem of time sharing, multitask coordination in these situations also requires mechanisms that can handle more complex scheduling processes, such as interruption and time-dependent resumption. Such complexity might seem to require a more general-purpose high-level controller that can carry out the appropriate scheduling and coordination functions when needed, across a wide variety of multitask situations. A recent model using the ACT-R architecture has been used to examine the functionality of a general purpose executive controller (Salvucci, 2005). The model proposes that generic multitasking abilities are accomplished through a goal-queuing mechanism that sets time-based priorities on the execution of different goals, and thus allows effective scheduling within the constraints of a serialized goal-execution process. The model was effectively applied to the task of driving in a virtual environment, with required control and monitoring of all its subtask components. It seems clear that this form of generic goal-queuing mechanism may represent the highest form of attentional control by specifying not only how attention gets allocated to a particular task, but also when and with what

priority the attentional allocation process occurs.

3.5. *Automaticity: Actions Without Attention?*

The general concept that behaviors executed repeatedly become less demanding and less effortful – a view that is clearly in line with subjective experience – has been studied at least since the dawn of modern psychology (James, 1890). More recently, this idea of automaticity has been considered to describe a specific mode of functioning in the mind/brain after extensive training in the execution of tasks. This automatic mode of processing enables performance to be qualitatively more efficient, robust, and rapid (Posner & Snyder, 1975; Schneider & Shiffrin, 1977). The key aspects that define automatic task processing are: (1) a decrease in effort, (2) an increase in speed along with practice, (3) no dependency on voluntary control, and (4) no interference with concurrent processes. A hybrid symbolic-connectionist computational architecture (CAP2) that accounts for these aspects can be found in Schneider and Chein (2003). It consists of a network of task-processing modules, each of which is a connectionist network linked with a central control system sending priority signals. The key process in this framework that enables a transition from controlled to automatic processing is a reduction in the requirement for such control signal intervention to ensure appropriate processing and selection in the distributed network of task-processing modules. Specifically, in the controlled processing state, control signal input is required for selection and amplification of the output of task-relevant processing modules, such that these outputs can be broadcast to other modules (e.g., those involved in response generation). In the automatic state, learning has occurred that enables certain outputs of a module to be coded as high priority, which then enables transmission to other modules, even in the absence of control system inputs. Thus, the transition from controlled to au-

tomatic processing in this model can be seen as a shift in whether attentional selection is governed by top-down or bottom-up biasing mechanisms.

One question that has been debated is whether automaticity reduces or even stops the demands of attention. The view taken in connectionist modeling is that automaticity does not completely shut off the requirement of attention, but simply reduces it. In other words, the role of controlled attention in task execution is not of the all-or-none variety. Graded and continuous attributes of automaticity, as well as of attention, should instead be considered. In the Stroop models of Cohen et al. (1990) described earlier, color naming is considered in need of attention. However, word reading, even though considered to be automatic, also still requires a top-down modulatory input from the task-demand units to generate a response within an appropriate timeframe. Therefore, word reading requires attentional control, although to a much lesser extent than color naming due to the stronger weights on the word-reading pathway.

A rather detailed analysis and model of the processes associated with automaticity can be found in the ACT-R framework (e.g., in Anderson, 1992). The basic view is that automaticity is due to the progressive compilation and associative linking of task-related production rules due to extensive training. The ACT-R framework has provided the most successful and comprehensive account to date regarding phenomena associated with automatization of processing, such as the power law rule of learning.

Another formal theory of automaticity can be found in Logan (2005). The key principle of this theory is that novel actions must be executed sequentially, according to a step-by-step algorithm. However, after the completion of each such action, a memory trace of its execution is formed. In the future, when that action is required again, it can be executed step by step as before or by accessing its memory, depending on which is faster. Each performance of an action and accumulation of experience leads

to the storage of further (discrete) instances of the action in memory, which in turn leads to a higher likelihood that one of these instances will be retrieved from memory and faster than the algorithm, thus producing automaticity.

3.6. *Unresolved Issues and Future Directions*

Several key issues remain unresolved in research on attentional control, and many research challenges still await a solution. A fundamental issue concerns the well-known capacity-limited nature of attention and cognitive control. Cognitive control is effortful, and the capacity to maintain task-relevant representations active even for intervals of seconds is a very limited ability, as several studies on these constraints have shown (Ansorge, 2004; Cowan, 2001; Engle, Kane, & Tuholski, 1999; Schneider & Shiffrin, 1977). But a clear theoretical justification for capacity constraints is still lacking, except for speculations that they are due to limitations of metabolic resources (Just et al., 2001) or that they are an emergent computational property arising from the necessity to constrain a massively parallel computer (the brain) into actions that have to be performed serially and unequivocally (Allport, 1989). A further speculation may have to do with competitive interactions between actively maintained goal representations in PFC, such that only a limited number can be sustained simultaneously without mutual interference or decay (O'Reilly, Braver, & Cohen, 1999; Usher & Cohen, 1999).

Another issue relates to exactly how attentional control is engaged and implemented and its relationship with conflict. The conflict-monitoring hypothesis of ACC–PFC interactions starts to tackle this issue, but convincing explanations and experimental verification of how conflict modulates control is still an open question. Is the information conveyed by the conflict signal precise enough to even address specific attentional control strategies that may be implemented in PFC? Other more general issues relate to the nature and function-

ing of the attention-related representations thought to be housed in PFC. These are usually referred to as rules, task demands, intentions, or goals, but explanations of how the anterior part of the neocortex implements and develops these representations is only just beginning. Such theoretical developments are critical for understanding the potentially specialized role of PFC representations in attention and for understanding their power in enabling flexible behavior.

One attempt to examine and understand the nature and development of PFC goal representations involved simulations training a single model to perform several different cognitive control tasks through an interleaved learning protocol (Rougier et al., 2005). As a result of this training, the model self-organized to develop abstract rule-like representations that preferentially coded dimensional properties of task stimuli. These representations were found to be sufficient to enable the model to successfully perform new attentional tasks, such as the Stroop without additional specialized training. Most importantly, the developed representations also enabled a high degree of within-task generalization, such that appropriate performance could be exhibited by the model for stimuli that it had never previously encountered during training. However, this model constitutes only a first attempt to understand the nature of PFC representations and their functionality. More complex forms of complex symbolic reasoning still remain to be addressed, for example, the dynamical recombination of different representations and how these interact with other cognitive systems. Nonetheless, it is an important manifestation of how computational modeling can provide an understanding of even hard dilemmas, such as flexible attentional control, without recourse to the homunculus.

Even setting aside questions of how goal-related representations develop, there are other important questions of the activation dynamics of such representations. For one, how is it possible to maintain a goal or intention for days and years, and not just seconds, as is usually modeled? These representations

cannot be explained just by active representations in PFC, but necessarily by some other flexible mechanism acting in a much larger time scale. One such mechanism may involve the storage and retrieval of goal information in episodic memory. However, the specifics of whether, how, and when such storage occurs are as yet unknown. A related issue concerns the scheduling of attentional control for goals and subgoals in the execution of complex tasks. The precise neural mechanisms involved in the coordination, transformation, and integration of stored and hierarchically organized information in complex task situations are still poorly understood. Recent empirical studies have begun to focus investigation on the most anterior part of the PFC as critical for a variety of goal-scheduling functions, such as branching (Koechlin et al., 1999), deferral (Burgess et al., 2003) and integration/coordination (Braver & Bongiolatti, 2002; De Pisapia & Braver, in press) during multitask conditions. But, as yet, there have been no computational models developed that can integrate and synthesize the accumulating data into an account of how anterior PFC mechanisms might specifically contribute to high-level multitasking functions.

A final important issue relates to how attention relates to other critical constructs and motivation, such as emotion, motivation, and consciousness. With regard to emotion, it is clear that any comprehensive theory of attention will need to address how attentional mechanisms are modulated by internal estimates of value. Yet, at this point, models of attention have been developed independently of affective/motivational considerations and vice versa. Nevertheless, the inclusion of neural mechanisms in attentional models that are also thought to have affective and motivational functions, such as the ACC and dopamine neurotransmitter system, may point to the route for these constructs to be eventually integrated within a unified framework.

With regard to consciousness, it seems critical to understand why voluntary attentional control and the effort it requires seem very prominent in subjective experience, whereas other forms of attentional modulation seem to go on in the absence of awareness. A recent review (Maia & Cleeremans, 2005) suggests the intriguing possibility that computational modeling of cognitive control and the biased competition framework could provide the theoretical path for an integration of attentional control with consciousness and working memory based on the idea of global competition between representations with the top-down biasing from PFC. These and other questions on attentional control and related cognitive constructs, as interesting as they are, remain without convincing answers.

4. Conclusion

This chapter has reviewed key computational models and theoretical directions pursued by researchers trying to understand the multifaceted phenomenon of attention. A broad division is drawn between theories and models addressing the mechanisms by which attention modulates specific aspects of perception (primarily visual) and those that have focused on goal-driven and task-oriented components of attention. Although the scope of the field is broad, the various accounts that have been put forth all seem to converge on the idea that attention can be understood as the mechanisms of focused selection and enhancement of currently processed information, and the suppression of perceived background aspects. Inquiring more specifically into how these mechanisms actually work has produced many more questions than answers, and this proliferation of unresolved issues likely will not end soon. On the other hand, over the last twenty years, there has been tremendous progress in the number and success of attempts to embody theoretical hypothesis into explicit computational and mathematical models. A particularly noteworthy point of convergence has been the widespread adoption of the biased competition framework as the core computational backbone of many attention models. More islands of

growing convergence will probably emerge in the coming years. Implemented models are the main instrument that researchers have available to substantiate or falsify their theories. The use of formal models that serve as explicit information-processing devices and that do not assume an internal observer or hidden homunculus will be critical in the effort to eventually fit, predict, and decompose human data from complex cognitive activities down to the most elemental components.

References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: The MIT Press.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

Anderson, J. R. (1992). Automaticity and the ACT* theory. *American Journal of Psychology*, *105*(2), 165–180.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *4*(111), 1036–1060.

Ansorge, U. (2004). Top-down contingencies of nonconscious priming revealed by dual-task interference. *The Quarterly Journal of Experimental Psychology: A, Human Experimental Psychology*, *57*(6), 1123–1148.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: Role of the subthalamic nucleus. *The Journal of Neuroscience*, *26*(9), 2424–2433.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.

Barch, D. M., Carter, C. S., Braver, T. S., Sabb, F. W., Noll, D. C., & Cohen, J. C. (1999). Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions. *Neuroimage*, *10*(6), 642–657.

Barch, D. M., Carter, C. S., & Cohen, J. D. (2004). Factors influencing Stroop performance in schizophrenia. *Neuropsychology*, *18*(3), 477–484.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Braun, J., Koch, C., & Davis, L. J. (2001). *Visual attention and cortical circuits*. Cambridge, MA: MIT Press.

Braver, T. S., Barch, D. M., & Cohen, J. D. (1999). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biological Psychiatry*, *46*, 312–328.

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage*, *15*(3), 523–536.

Braver T. S., & Cohen J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII; control of cognitive processes* (pp.713–737). Cambridge, MA: MIT Press.

Braver, T. S., & Cohen, J. D. (2001). Working memory, cognitive control, and the prefrontal cortex: Computational and empirical studies. *Cognitive Processing*, *2*, 25–55.

Braver, T. S., Cohen, J. D., & Barch, D. M. (2002). The role of the prefrontal cortex in normal and disordered cognitive control: A cognitive neuroscience perspective. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 428–448). Oxford, UK: Oxford University Press.

Braver, T. S., Cohen, J. D., & Servan-Schreiber, D. (1995). Neural network simulations of schizophrenic performance in a variant of the CPT-AX: A predicted double dissociation. *Schizophrenia Research*, *15*(1–2), 110.

Broadbent, E. D. (1958). *Perception and communication*. London: Pergamon.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118–1121.

Brown, J. W., Reynolds, J. R., & Braver, T. S. (2007). A computational model of fractionated conflict-control mechanisms in task-switching. *Cognitive Psychology*, *55*, 37–85.

Brunel, N., & Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of

object working memory dominated by recurrent inhibition. *Journal of Computational Neuroscience, 11*, 63–85.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523–547.

Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 353*(1373), 1271–1281.

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112*(2), 291–328.

Burgess, P. W., Scott, S. K., & Frith, C. D. (2003). The role of the rostral frontal cortex (area 10) in prospective memory: A lateral versus medial dissociation. *Neuropsychologia 41*(8), 906–918.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108*(4), 847–869.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747–749.

Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research, 62*(2–3), 182–194.

Chang, L., Speck, O., Miller, E. N., Braun, J., Jovicich, J., Koch, C. et al. (2001). Neural correlates of attention and working memory deficits in HIV patients. *Neurology, 57*(6), 1001–1007.

Cohen, J. D., Braver, T. S., & Brown, J. W. (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion in Neurobiology 12*, 223–229.

Cohen, J. D., Braver, T. S., & O'Reilly, R. C. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 351*(1346), 1515–1527.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332–361.

Cohen, J. D., & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà & M. Moscovitch (Eds.), *Attention and perfor-*

mance *XV* (pp. 1–19). Cambridge, MA: MIT Press.

Cohen, J. D., Romero, R. D., Farah, M. J., & Servan-Schreiber, D. (1994). Mechanisms of spatial attention: The relation of macrostructure to microstructure in parietal neglect. *Journal of Cognitive Neuroscience, 6*(4), 377–387.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychological Review, 99*, 45–77.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87–185.

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing, 69*(10–12), 1322–1326.

De Pisapia, N., & Braver, T. S. (2007). Functional specializations in lateral prefrontal cortex associated with the integration and segregation of information in working memory. *Cerebral Cortex, 17*, 993–1006.

Deco, G. (2001). Biased competition mechanisms for visual attention in a multimodular neurodynamical system. In S. Wermter, J. Austin, D. Willshaw, and M. Elshaw (eds.), *Emergent neural computational architectures based on neuroscience: Towards neuroscience-inspired computing* (pp. 114–126). Berlin: Springer.

Deco, G., & Lee, T. S. (2004). The role of early visual cortex in visual integration: A neural model of recurrent interaction. *The European Journal of Neuroscience, 20*(4), 1089–1100.

Deco, G., Pollatos, O., & Zihl, J. (2002). The time course of selective visual attention: Theory and experiments. *Vision Research, 42*(27), 2925–2945.

Deco, G., & Rolls, E. T. (2002). Object-based visual neglect: A computational hypothesis. *The European Journal of Neuroscience, 16*(10), 1994–2000.

Deco, G., & Rolls, E. T. (2003). Attention and working memory: A dynamical model of neuronal activity in the prefrontal cortex. *The European Journal of Neuroscience, 18*(8), 2374–2390.

Deco, G., & Rolls, E. T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research, 44*(6), 621–642.

Deco, G., & Rolls, E. T. (2005a). Attention, short-term memory, and action selection: A unifying theory. *Progress in Neurobiology, 76*(4), 236–256.

Deco, G., & Rolls, E. T. (2005b). Neurodynamics of biased competition and cooperation for attention: A model with spiking neurons. *Journal of Neurophysiology, 94*(1), 295–313.

Deco, G., Rolls, E. T., & Horwitz, B. (2004). "What" and "where" in visual working memory: A computational neurodynamical perspective for integrating FMRI and single-neuron data. *Journal of Cognitive Neuroscience, 16*(4), 683–701.

Deco, G., & Zihl, J. (2001a). A neurodynamical model of visual attention: Feedback enhancement of spatial resolution in a hierarchical system. *Journal of Computational Neuroscience, 10*(3), 231–253.

Deco, G., & Zihl, J. (2001b). Top-down selective visual attention: A neurodynamical approach. *Visual Cognition, 8*(1), 119–140.

Deco, G., & Zihl, J. (2004). A biased competition based neurodynamical model of visual neglect. *Medical Engineering & Physics, 26*(9), 733–743.

Dehaene, S., & Changeux, J. P. (1991). The Wisconsin card sorting test: Theoretical analysis and modeling in a neuronal network. *Cerebral Cortex, 1*(1), 62–79.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Durstewitz, D., Kelc, M., & Gunturkun, O. (1999). A neurocomputational theory of the dopaminergic modulation of working memory functions. *Journal of Neuroscience, 19*(7), 2807–2822.

Durstewitz, D., Seamans, J. K., & Sejnowski, T. J. (2000). Neurocomputational models of working memory. *Nature Neuroscience, 3*(Suppl) 1184–1191.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 102–134). New York: Cambridge University Press.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44*(3), 297–337.

Hamker, F. H. (2003). The reentry hypothesis: Linking eye movements to visual perception. *Journal of Vision, 3*(11), 808–816.

Heinke, D., Deco, G., Zihl, J., & Humphreys, G. (2002). A computational neuroscience account of visual neglect. *Neurocomputing, 44–46*, 811–816.

Heinke, D., & Humphreys, G. W. (2003). Attention, spatial representation, and visual neglect: Simulating emergent attention and spatial memory in the selective attention for identification model (SAIM). *Psychological Review, 110*(1), 29–87.

Heinke, D., & Humphreys, G. W. (2005). Computational models of visual selective attention: A review. In G. Houghton (Ed.), *Connectionist models in psychology* (pp. 273–312). London: Psychology Press.

Herd, S. A., Banich, M. T., & O'Reilly, R. C. (2006). Neural mechanisms of cognitive control: An integrative model of stroop task performance and FMRI data. *Journal of Cognitive Neuroscience, 18*(1), 22–32.

Houghton, G., & Tipper, S. P. (1996). Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. *Brain and Cognition, 30*(1), 20–43.

Humphreys, G. W., & Muller, H. J. (1993). SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25*(1), 43–110.

Itti, L., & Baldi, P. (2005). A principled approach to detecting surprising events in video. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 1*, 631–637.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature reviews. Neuroscience, 2*(3), 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, 20*(11), 1254–1259.

James, W. (1890). *Principle of Psychology.* Dover Publications 1950, vol. 1.

Just, M. A., Carpenter, P. A., Keller, T. A., Emery, L., Zajac, H., & Thulborn, K. R. (2001). Interdependence of nonoverlapping

cortical systems in dual cognitive tasks. *Neuroimage, 14*(2), 417–426.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219–227.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature, 399*, 148–151.

Kwak, H. W., & Egeth, H. (1992). Consequences of allocating attention to locations and to other attributes. *Perception & Psychophysics, 51*(5), 455–464.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience, 2*(4), 375–381.

Lee, T. S. (1996). Image representation using 2D gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(10), 959–971.

Logan, G. D. (2005). The time it takes to switch attention. *Psychonomic Bulletin & Review, 12*(4), 647–653.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*(2), 393–434.

Lovett, M. C. (2002). Modeling selective attention: Not just another model of Stroop. *Cognitive Systems Research, 3*(1), 67–76.

Maia, T. V., & Cleeremans, A. (2005). Consciousness: Converging insights from connectionist modeling and neuroscience. *Trends in Cognitive Sciences, 9*(8), 397–404.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86*, 287–330.

McClelland, J. L., & Rumelhart, D. E. (1981). An interative activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375–407.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110*, 422–471.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104*, 3–65.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 21*, 167–202.

Mozer, M. C., Shettel, M., & Vecera, S. (2006). Top-down control of visual attention: A rational account. In Y. Weiss, B. Schoelkopf, & J. Platt (Eds.), *Neural information processing systems* (pp. 923–930). Cambridge, MA: MIT Press.

Mozer, M. C., & Sitton, M. (1998). Computational modeling of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341–393). London: Psychology Press.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*(2), 205–231.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and Self-regulation* (Vol. 4, pp. 1–18). New York: Plenum Press.

Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience, 13*(11), 4700–4719.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science, 314*, 91–94.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically-based computational model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the ming by simulating the brain.* Cambridge, MA: MIT Press.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107–123.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*(2), 220–244.

Phaf, R. H., Van der Heijden, A. H., & Hudson, P. T. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22*(3), 273–341.

Posner, M. I., Cohen, Y., & Rafal, R. D. (1982). Neural systems control of spatial orienting.

Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 298(1089), 187–198.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Reynolds, J. R., Braver, T. S., Brown, J. W., & Stigchel, S. (2006). Computational and neural mechanisms of task-switching. *Neurocomputing, 69,* 1332–1336.

Roelofs, A. (2000). Control of language: A computational account of the Stroop asymmetry. In *Proceedings of the Third International Conference on Cognitive Modeling.* Veenendaal, The Netherlands: Universal Press.

Rougier, N. P., Noelle, D., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005). Prefrontal cortex and the flexibility of cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences, 102*(20), 7338–7343.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1 and 2). Cambridge, MA: MIT Press.

Salvucci, D. D. (2005). A multitasking general executive for compound continuous tasks. *Cognitive Science, 29,* 457–492.

Schneider, D. W., & Logan, G. D. (2005). Modeling task switching without switching tasks: A short-term priming account of explicitly cued performance. *Journal of Experimental Psychology: General, 34*(3), 343–367.

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science, 27*(3), 525–559.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Servan-Schreiber, D., Bruno, R., Carter, C., & Cohen, J. (1998). Dopamine and the mechanisms of cognition: Part I. A neural network model predicting dopamine effects on selective attention. *Biological Psychiatry, 43*(10), 713–722.

Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences, 8*(5), 223–230.

Sohn, M.-H., & Anderson, J. R. (2003). Stimulus-related priming during task switching. *Memory & Cognition, 31*(5), 775–780.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Treisman, A. (1999). Feature binding, attention and object perception. In J. W. Humphreys & J. Duncan (Eds.), *Attention space and action* (pp. 91–111). Oxford, UK: Oxford University Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Tsotsos, K. J., Culhane, M. S., Wai, K. W. Y., Lai, Y., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78,* 507–545.

Usher, M., & Cohen, J. D. (1999). Short term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphries, & A. Olsen (Eds.), *Connectionist models in cognitive neuroscience* (pp. 78–91). The 5th Neural Computation and Psychology Workshop. Springer Verlag, University of Birmingham, UK.

van de Laar, P., Heskes, T., & Gielen, S. (1997). Task-dependent learning of attention. *Neural Networks, 10*(6), 981–992.

van der Heijden, A. H., & Bem, S. (1997). Successive approximations to an adequate model of attention. *Consciousness and Cognition, 6*(2–3), 413–428.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron, 46,* 681–692.

Ward, R. (1999). Interaction between perception and action systems: A model for selective action. In W. G. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action – Studies in cognitive neuroscience* (pp. 311–332). Oxford, UK: Oxford University Press.

Wolfe, C. D., & Bell, M. A. (2004). Working memory and inhibitory control in early childhood: Contributions from physiology, temperament, and language. *Developmental Psychobiology, 44*(1), 68–83.

Wolfe, J. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238.