2

A Model of Inhibitory Mechanisms in Selective Attention

George Houghton and Steven P. Tipper

I. INTRODUCTION

The complex environments encountered and internally represented by many organisms contain numerous objects toward which action could be directed, or which the organism may need to avoid. Given the simultaneous presence of many objects of different adaptive value, and affording a variety of different responses, it is extremely important for coherent, organized behavior that actions be selectively directed toward one object at a time. In this chapter we propose that the means by which internal goal states mediate the interaction between perception and action is the mechanism of selective attention.

For a complete understanding of selective attention, investigations have to take place at a number of different levels (Marr, 1982). For example, the precise adaptive function of attention must be specified within the context of an organism's normal interactions with the world, as well as the kinds of representations on which attentional processes act. Algorithms for executing attentional processes need to be specified and tested experimentally. Experimental evidence may be drawn from the cognitive level, where performance by human subjects is carefully analyzed in laboratory tasks (Broadbent, 1958; Kahneman, 1973; Posner, 1978), and from the neurophysiological level, where the activity of single cells, or an animal's performance after lesions, can be assessed during attentional tasks (Moran & Desimone, 1985; Picton, Stuss, & Marshall, 1986; Posner & Driver, 1992).

The interaction between the cognitive and neurophysiological levels of investigation provides mutual constraints for models of attention. However, most models derived from experimental results are informal (e.g., Tipper, 1985), and hence it is often difficult to detect inconsistencies in logic, or to demonstrate that the informal theory is sufficiently precisely specified to constitute an effective explanation of the data under consideration. Such models therefore need to be made maximally explicit, for instance by being specified mathematically and instantiated in computer programs which can work out the predictions of the model in detail. In this chapter we develop a preliminary neural network model of certain aspects of voluntary selective attention which is consistent with our views regarding the central purpose ("why") and method ("how") of selective attention.

Section II discusses our overall approach to selective attention (our "computational theory," using Marr's (1982) term) and considers a variety of data in support of it. In Section III we develop a neural network model that formalizes certain aspects of the theory presented in Section II, and in Section IV we show how the model accounts for a variety of existing data.

II. THEORIES OF ATTENTION

In discussing theories of selective attention, and the place of our theory in relation to others, it is useful to identify two dimensions of variation: the locus of selection in the pathway from perception to response (e.g., late vs. early), and the kind of mechanism used in selection (e.g., amplificatory, "the spotlight," or inhibitory or both). We first consider the issue of the locus of selection, which we believe to be central to the issue of the functional role of attention, and then consider the nature of the mechanisms involved.

A. Locus and Function of Selective Attention

With regard to the issue of the locus of selection, two positions have emerged which we refer to as the *precategorical* and *postcategorical* positions. The precategorical view proposes that attention is critical for perceptual processes, that without attention only low levels of analysis can take place, for instance detecting edges and motion (Triesman & Gelade, 1980). More specifically, it proposes that attention is necessary for the categorization of objects, hence the term precategorical. The postcategorical position proposes that perceptual processes, including grouping and (implicit) categorization, are largely automatic, at least for familiar stimuli. Selection acts on objects (or perceptual groups) for the purposes of controlling action towards them.

The models of Koch and Ullman (1985) and LaBerge and Brown (1989) are of this type. For example, LaBerge and Brown's model describes three properties of selective attention. First, selection takes place at early levels of perceptual input, allowing one object to enter a limited capacity identification system; second, attentional modulation of low-level stimulus attributes is carried out by reference to their location; and third, selection is based on excitation, where the perceptual features of a stimulus receive facilitated transmission into the identification system. In contrast, the model we describe is qualitatively different on all of these dimensions. Our model proposes that selection can take place after perceptual grouping and semantic analysis, that attention accesses object-based representations, and that a central mechanism in the selection process is active inhibition of distracting information.

The position we take in this chapter is based on the computational level of analysis emphasized by Marr (1982). Such an analysis specifies the functions for which a particular system has evolved, considering both biological goals and the environment in which such goals are to be achieved. In other words, attention must be considered in the global context of an organism's self-organization of its behavior, in pursuit of its socially and biologically acquired goals (Luria, 1973; Norman & Shallice, 1986). Thus we suggest that in the context of an organism's encounters with its familiar, natural environments, the role of attention in low-level perceptual analysis may be quite limited. Schema-based expectations may facilitate the perception and comprehension of a scene without the need to focus attention serially on each object in the scene (Biederman, Mezzanotte, & Rabinowitz, 1981; Palmer, 1975; Friedman, 1979; Treisman, 1986). Indeed, if attention is necessary for object perception, it is difficult to see how it can be moved from object to object in the first place. Furthermore, the constraints provided by information distributed throughout a scene (such as objects appearing in conjunction with other objects, being supported by them, having the ability to occlude other objects, having a typical locus in the scene, etc.) reflect the identity of an object and take part in the construction of its internal representation. These "normal" conditions contrast with the highly nonpredictable decontextualized situations encountered in the typical experimental study of attention.

It is thus likely that highly efficient parallel perceptual analysis takes place in familiar situations. As Neumann (1987) noted, the brain processes an immense amount of perceptual information (from many modalities) in parallel, and thus information handling limitations in the nervous system may not be critical at the level of perceptual analysis, but rather at more central levels, specifically in the linking of perception and action, and the perceptual control of effector systems: one hand cannot raise two beer glasses to the mouth for drinking, just as one visual object naming system cannot name two objects at once.

Apart from such general considerations, experiments with the negative priming paradigm (Tipper, 1985; discussed in more detail later) have shown clear *semantic* effects of distractors. In these priming studies, an ignored picture of a cat can affect the subsequent processing of a semantically related word such as *dog* (Tipper & Driver, 1988). Note that there is no physical resemblance between the picture and the word, so any interactions between them must be at semantic levels. Other work has demonstrated that the semantic properties of ignored objects can influence the concurrent processing of a target stimulus (Stroop, 1935). In the priming studies of Tipper and his colleagues, the distracting stimuli are usually presented very briefly and subjects are typically unable to report awareness of the objects' identity. It is argued from such evidence that the prime display is analyzed in parallel by the visual system to semantic levels, but that the selective attention mechanism only allows information from an object possessing the target property into the response systems (and into awareness). Information relating to the distractor is inhibited and this in some way retards later responses to semantically related items (this inhibitory element will be discussed further).

It may be objected here that it is pointless to construct representations of multiple objects in parallel, only to have to suppress information from most of them for the purposes of generating coherent thought and action. We would argue, however, that the identification of objects is facilitated by their occurring in familiar contexts with other known objects (Biederman, 1972; Biederman, Mezzanotte, & Rabinowitz, 1982). In addition, we suggest that interactions with complex real world environments *require* object-based representations of distractor objects, rather than free-floating features. Consider reaching for potato chips from a bag placed in the middle of a table. Also on the table are a variety of bottles, glasses, ashtrays, and so on that one must reach around or over. Successful behavior could not be achieved if only the target stimulus was fully represented and distractors were only analyzed in terms of low-level features. Rather, we propose that to achieve the behavioral goal, the accommodation of a detailed form of an action to the nontarget objects is required, such as when reaching around or over obstacles. Such indirect actions can only be guided by object- or action-centered representations (see Tipper, Weaver, & Houghton, in press). Thus human perceptual systems have evolved to produce internal representations of both objects forming the focus of current goal-directed behavior and those providing the context.

We therefore propose that the role of attention in overall behavior lies not so much in the creation of an internal representation of the environment, but rather in the linking of the appropriate action with the appropriate object in contexts which may afford an arbitrary number of such linkages, the great majority of which, at any time, will be disruptive to the organism's goal-seeking behavior (Luria, 1973). In the model presented here, we suppose processing of visual inputs to be taking place in parallel, up to categorical levels, without the intervention of attentional processes, and that selection involves the "binding" of information in target object representations (e.g., location, form, etc.) to variables associated with response schemas (Arbib, Iberall, & Lyons, 1987; Arbib, 1990). Object representations *compete* for control of action by binding the parameters of action schemas, and selective attention acts to modulate this competition in favor of target objects.[1]

B. Amplification and Inhibition in Selective Attention

Whatever the location and function of selection, we can consider two mechanisms by which it may be achieved: amplification and inhibition (enhancement and suppression, respectively). Spotlight models of attention are based on amplification and argue that selection of a target from a distractor is primarily an excitatory process. Attention is directed toward some part of the input field, like the beam of a spotlight, and information within the beam is amplified, allowing it to receive further processing beyond that of initial preattentive analysis. The initial representations activated by ignored inputs passively decay back to resting levels (Broadbent, 1970; Van der Heijden, 1981). Such a model appears to suggest no role for inhibitory mechanisms. An alternative account, one which we develop here, argues that the remarkably efficient selection

[1] The adaptive value of the capacity to select out one object from an array of potential targets is perhaps most clearly illustrated in the phenomenon of predatory attack (see Curio, 1976). In terms of our model, this requires that the typical predator select one target prey and bind the parameters of the attack response to the movements of the target. Many prey animals move together in groups (e.g., schools of fish) and it has been found that the attack response of predators becomes disorganized and less successful when confronted with such groups (as opposed to isolated prey). Hunting success has been found to vary inversely with school size for some predatory fish (Neill & Cullen, 1974) and the reason for this seems to be that the copresence of multiple, highly similar potential targets disrupts the selective coupling of the attack response. As response to this, some predators have evolved the tactic of breaking up fish schools, thereby increasing the number of isolated stray individuals, which are then selected for attack. We would hypothesize that this is because the absence of distractors near strays facilitates the sustained selection necessary for pursuit and capture. It is thus possible that schooling behavior has evolved in part to exploit limitations in the attentional systems of predators.

found in experimental tasks may be better accounted for by a dual mechanism consisting of the previously described excitatory component, supplemented by an inhibitory component that acts to suppress competing information derived from the analysis of the distractor. We show in the model how such combined processes of excitation and inhibition can rapidly differentiate the internal representations of the target and distractor objects (Tipper, 1985).

The point may be made that postulating a dual mechanism is unparsimonious, if all one is ultimately doing is separating out information-bearing signals by selective gain control—all that is needed for this is either an amplificatory mechanism or an inhibitory one, but not both. We would argue, however, that a dual (opponent) mechanism is advantageous for at least two reasons. First, any gain control mechanism implemented in biological (i.e., neural) hardware can only operate within finite limits, and thus the rate at which one signal can be boosted with respect to another (remaining constant) must have some finite upper bound. A dual mechanism, acting in parallel, can boost a target signal while suppressing a distractor, thereby effectively doubling the rate at which target and distractor (signal and noise) can be pulled apart. The second reason is perhaps less obvious. Any signals in a biological information processing system must have a limited dynamic range—maximum and minimum amplitudes. Suppose we arbitrarily scale this range in the interval [0,1], with 0 representing the floor and 1 the ceiling. It may happen that this system is simultaneously presented with two signals, T and D, and that both are at high levels, say 0.9 on our scale. Suppose that one signal, T, is to be "attended." In this case a "spotlight" mechanism can induce a maximum signal amplification of about 12% ($T = 1.0$, $D = 0.9$), at which point the target signal T saturates. This gap is unlikely to be sufficient to prevent debilitating interference from D, which remains at high amplitude. Clearly what is needed in such a case is some ability to suppress D to a significant extent. The converse argument holds for low amplitude signals, where selective amplification would be more effective than inhibition of low amplitude distractors. A system with both components would thus be maximally effective over the entire dynamic range, exhibiting more rapid selection of targets and better performance in situations of low signal-to-noise ratios.

The emphasis we place here on the ability to inhibit distracting information is not new. According to Wundt (1904),

> The basic phenomenon of all intellectual achievement is the so-called concentration of attention. It is understandable that in the appraisal of this phenomenon we attach importance first and therefore too exclusively to its positive side, to the grasping and clarification of certain presentations. But for the physiological appraisal it is clear that it is the negative side, the inhibition of the inflow of all other disturbing excitations . . . which is more important. (p. 481)

From this perspective, the excitation-spotlight metaphor is perhaps better regarded as representing the "view from consciousness," that is, it is how things appear to the conscious mind, which deals only with the results of selection in the control of thought and action. The actual mechanisms of selection, and the levels of preattentive analysis, may bear little resemblance to what is suggested by such an introspective viewpoint. We now consider some of the experimental evidence for the importance of late, inhibitory processes in selective attention.

Evidence for inhibitory mechanisms in selective attention comes from both physiological and cognitive research. Within the neurophysiological paradigm, Moran and Desimone (1985) demonstrated that when a stimulus was ignored while attention was directed to another stimulus, the response of the cell in whose receptive field the ignored stimulus was located was suppressed. They suggested that suppression of distractors may be the primary mechanism of selection in these cortical regions (V4 and the inferior temporal lobe). Electrophysiological techniques have also provided evidence for the inhibition of distractors. Arsten et al. (1983) demonstrated that naloxone improves selectivity by increasing the suppression of distractors in the frontal lobe, rather than by changing the analysis of the attended targets. Much neuropsychological evidence points to the role of areas of the prefrontal lobes in the direction and maintenance of attention, and to a strong inhibitory component in this function. For instance, on reviewing an extensive body of evidence on the subject, Fuster (1980) concludes:

> [N]europsychological evidence . . . points to the essential role of the orbitomedial prefrontal cortex in the suppression and control of [sources of] interference . . . [T]hat role may be considered inhibitory and part of . . . the selective attention that the animal must direct and maintain for the proper conduct of the behavioral sequence. (p. 187)

Fuster further remarks that the mechanisms by which prefrontal attentional systems interact with other cortical areas (e.g., posterior sensory areas) are unknown. It seems likely that progress in this area will depend not only on further physiological investigation, but also on the development of much more detailed models of selective attention that can be interpreted physiologically.

Much of the cognitive evidence for inhibition of distractors arises from the negative priming paradigm (Dalrymple-Alford & Budayr, 1966; Neill, 1977; Tipper, 1985). If the internal representations of a to-be-ignored object are associated with inhibition during selection and execution of the responses to the target object, the processing of a subsequent stimulus requiring the inhibited representations should be impaired. Therefore, in a priming procedure, when an ignored stimulus in a prime trial is subsequently presented as a probe for rapid identification, reaction time (RT) to name this probe should be increased.

In the procedure illustrated in Figure 1 for example, subjects were instructed to attend to the red object (solid line) and ignore the green object (broken line). In the ignored repetition condition the ignored prime (green object) reappeared as the red object in the probe. As predicted by the inhibition hypothesis, RTs were longer in this situation than in the control condition where there was no relationship between the objects in the prime and probe displays. Such a result is consistently observed in negative priming tasks, and the evidence for the active inhibition of distractors thus obtained is now quite substantial and has been observed with a wide variety of experimental stimuli. These stimuli include:

1. Words (Beech, Baylis, Tipper, McManus, & Agar, 1991; Hoffman & MacMillan, 1985; Fuentes & Tudela, 1992; Hung & Tzeng, 1989; Tipper & Driver, 1988; Yee, 1991)

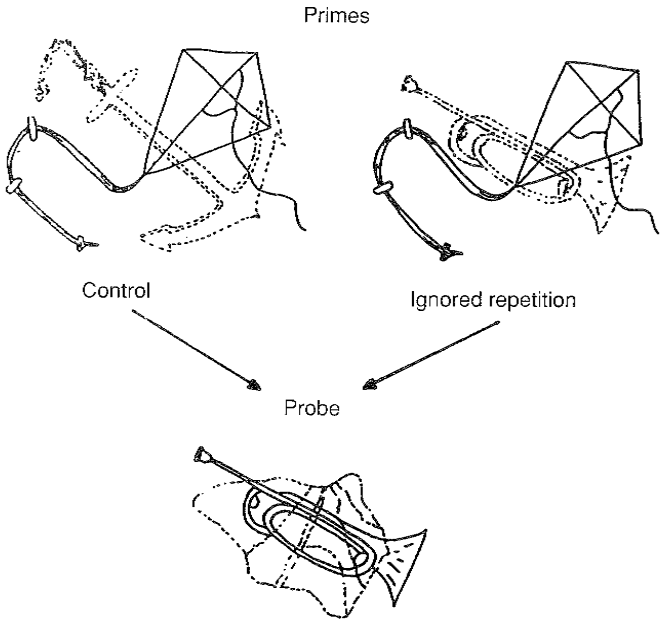

Figure 1
Sample displays from a study demonstrating negative priming. Subjects were required to name the red object (solid line) while ignoring a green distractor (broken line). In the control prime display, the target and distractor were unrelated to the subsequent probe, but in the ignored repetition condition, the ignored prime was the same as the subsequent probe target. Negative priming was revealed by longer reaction times to name the probe after presentation of the ignored repetition prime than after presentation of the control prime. The probe target was superimposed over a neutral, meaningless distractor.

2. Stroop color words (Beech, Agar, & Baylis, 1989; Beech, Baylis, Smithson, & Claridge, 1989; Beech & Claridge, 1987; Beech, Powell, McWilliams & Claridge, 1989, 1990; Benoit et al., 1992; Dalrymple-Alford & Budayr, 1966; Enright & Beech, 1990; Lowe, 1979, 1985; McLaren & Bryson, 1988; Neill, 1977; Neill & Westberry, 1987; Tipper, Bourque, Anderson, & Brehaut, 1989)
3. Local-global letters (Briand, in preparation; Baylis & Tipper, unpublished)
4. Letters (Hasher, Stoltzfus, Zacks, & Rympa, 1991; McDowd & Oseas-Kreger, 1991; Neumann & DeSchepper, 1991, 1992; Tipper & Cranston, 1985; Tipper, MacQueen, & Brehaut, 1988)
5. Pictures (Gernsbacher & Faust, 1991; McLaren, 1989; Tipper, 1985; Tipper & Driver, 1988; Tipper, Weaver, Cameron, Brehaut, & Bastedo, 1991)
6. Color (DeSchepper, Khurana, O'Connell, & Wilson, in preparation)
7. Random shapes (DeSchepper & Treisman, 1991)

Furthermore, negative priming has been obtained when report of the target locus is required, as opposed to target identification tasks (Tipper, Brehaut, & Driver, 1990; Tipper & McLaren, 1990; Tipper, Weaver, Kirkpatrick, & Lewis, 1991), and when the target and distractor stimuli are presented in different perceptual modalities (Driver & Baylis, 1993; Greenwald, 1972). The effects have also been observed when subjects engage in real-world tasks such as reaching towards stimuli (Tipper, Lortie, & Baylis, 1992). Similar patterns of results in the latter procedure are observed in both humans and infrahuman primates (Taffe, Moore, Tipper, & Baylis, 1991).

A central concern in this work has been the investigation of the locus of the inhibitory processes in the pathways from stimulus to response. As previously described, the evidence is clear that inhibition can act on semantic (categorical) level representations (indeed, the finding of negative priming effects between semantic associates is used as evidence that distractors are processed to semantic levels without necessarily reaching awareness). Other studies, in which subjects make different types of responses to prime and probe trials (Tipper et al., 1988), indicate that the inhibitory effects are not associated with specific motor responses. We might conjecture that inhibition is associated with those aspects of the stimulus most relevant to the goals of the organism. In tasks involving naming responses, this is semantic identity. However, when analysis of color (DeSchepper et al., in preparation) or the production of specific reaching responses (Tipper et al., in press) are central components of the task, then inhibition can be associated with the perceptual property in the former case, and with action-centered representations in the latter case. Thus, the inhibitory effects may not be associated with one particular level of internal representation, but might act flexibly to coordinate the link between perception and action at whatever level is optimal given the current task and stimulus situation. One important consequence of this position is that, because organisms typically interact with objects, it is most parsimonious for attention to access object-based representations (rather than, say, simple spatial coordinates; see Duncan, 1980, for elaboration of this point). Experimental evidence has, indeed, supported such a contention. The inhibitory component of the selection mechanism is associated with objects, so that as objects move through space, inhibition can move with them—that is, it is not tied to a location (Tipper et al., 1990; Tipper, Driver, & Weaver, 1991).

The proposed function of inhibitory mechanisms in our conception is that they assist in the efficient foregrounding of target information and reduce interference from competing distractors. If this is so, and if the degree of negative priming found is an indication of the strength of the inhibitory process, then we would predict an inverse relationship between interference and negative priming: less negative priming should indicate weaker inhibitory processes and hence greater interference. Evidence for such a relationship has been observed from a number of sources which have investigated individual differences in selective attention. It is well established that some populations have particular difficulty responding to stimuli when distractors are present, appearing to be unable to efficiently select. If inhibition is a mechanism of selection, then it may be predicted that such populations have less effective inhibitory mechanisms and will show less negative priming. Experiments investigating a variety of populations known to have selection difficulties support this suggestion. Thus, reduced negative priming has been observed in the following populations:

1. Children (Tipper et al., 1989)
2. Children with attentional deficit disorder (McLaren, 1989)
3. The elderly (Hasher et al., 1991; McDowd & Oseas-Kreger, 1991; Tipper, 1991)
4. Obsessionals (Enright & Beech, 1990)
5. Subjects who report high cognitive failure (Tipper & Baylis, 1987)
6. Schizophrenics (Beech et al., 1989)
7. High schizotypes (Beech & Claridge, 1987)
8. Depressed patients (Benoit et al., 1992)
9. Alzheimer's patients (Mueller & Baylis, in preparation)

Similarly, Gernsbacher & Faust (1991, Experiment 3) showed that subjects who scored low on story comprehension tests seem to be less able to suppress to-be-ignored stimuli than were control subjects (as indexed by the degree of interference caused by to-be-ignored stimuli on a subsequent response).

In the opposite direction, Beech et al. (1990) have shown that chlorpromazine (a dopamine blocking neuroleptic that reduces the symptoms of schizophrenia) increases negative priming. The results related to schizophrenia are of particular interest in that some recent models of schizophrenia suggest that one element of the schizophrenic syndrome involves loss of selectivity due to weakened inhibitory filtering systems. For instance, Swerdlow and Koob (1987) suggest that in schizophrenia, loss of inhibitory processes means that "appropriate filtering and amplification of cortical information cannot occur . . . and irrelevant and relevant cognitive or emotional activity are not segregated" (p. 204). The notion of "relevance" must surely be defined in relation to the subject's goal or drive states. In a similar vein, Gray, Feldon, Rawlins, Hemsley, and Smith (1991) postulate that the schizophrenic syndrome involves "the weakening of inhibitory processes crucial to conscious attention," which allows "the intrusion into awareness of aspects of the environment not normally perceived" (Gray et al., 1991, p. 2). The authors further suggest that in normals, selection is for goal-relevant stimuli, and therefore schizophrenics suffer a deficit in the ability to filter incoming (or endogenously generated) information on the basis of goal-derived targets. Further investigation of this intriguing area would clearly benefit from the development of explicit models of inhibitory filtering in normal subjects. We hope the following model will make a useful contribution.

Thus far we have discussed only evidence for inhibitory processes being associated with nontarget (distractor) items. There is evidence from other sources that attended items become actively suppressed when attention is switched. This is illustrated, for example, by the phenomenon of inhibition of return (Posner & Cohen, 1984), whereby responses to stimuli at previously attended locations are retarded, following a switch of attention, compared to items at new (never-attended) locations. This suggests that as a spatial attentional target is switched, the shift of attention involves inhibition of the last attended location. Similar results have been found involving conceptual entities and objects, rather than simply locations. For instance, in work on text comprehension, it has been found that when a text requires the switching of attentional focus from one protagonist to another, responses related to the previous focus become retarded (Gernsbacher, 1989). Within the attentional literature per se, evidence for object-centered inhibition of return is reported by Tipper et al. (1991), who show that when attention is cued to a *moving* object and then removed from it, a subsequent response made to that object is retarded relative to a response made to a different object. Because the objects are moving continuously, the inhibitory effect cannot be tied to a location, but instead moves with the object. These studies show that inhibition can be associated not only with distractor objects, but also with conceptual entities which have been previously attended. The work on text comprehension indicates that the inhibition of attended items occurs at the point at which the last attended item becomes a potential distractor, that is, when a shift of focus is required. It is of some interest to our work to consider whether these disparate inhibitory effects, involving both attended and unattended items, might arise from a single underlying mechanism. This issue has been raised before (see, e.g., Tipper et al., 1991; Gernsbacher & Faust, 1991), but thus far no explicit candidate mechanisms have been proposed.

This concludes the informal overview of our theory of attention and its current empirical basis. Thus far no precise mechanism has been proposed to account for the negative priming effects, or the active suppression of attended items, beyond general statements concerning inhibition becoming associated with internal representations and processes (Tipper, 1985; Gernsbacher & Faust, 1991). Proposals made in the experimental literature fail to provide answers to numerous questions, such as: What is the source of the inhibitory processes? How do they locate their targets (i.e., know relevant from irrelevant information)? How are they are initiated? How are they terminated? What is their timecourse more generally? The following model attempts to remedy this situation by proposing a detailed model of the dynamics of selective attention. The basic premises of the model, that selection is important in behavioral control, that it is object based, and so on, are derived from the extensive experimental literature just described. These premises are supplemented by a formalized selection mechanism which accounts in detail for a variety of previously obtained results. The proposed mechanism also motivates new experiments and contains nonobvious properties which we believe may contribute to the integration of different areas of research.

III. A MODEL OF THE DYNAMICS OF SELECTIVE ATTENTION

The previous discussion outlines the basic conception of selective attention, which the model described in this section attempts to develop more formally. We see the purpose of our modeling efforts as being threefold:

1. Increased theoretical precision. We believe the previously cited data provide ample evidence for inhibitory selection mechanisms acting on central (object-based) representations. However, current proposals regarding the nature of the mechanisms remain vague and intuitive. Thus ongoing empirical investigations need to be linked to a parallel program of theoretical development if a detailed understanding of the relevant phenomena is to be attained. We intend the model presented here to be a first step toward a precise mechanistic account of central selection processes, as employed in a wide variety of perceptual and cognitive tasks.
2. Increased predictive power. As well as providing more detailed accounts of known results, precise models can be expected to generate predictions that cannot be derived from intuition-bound formulations.
3. Theoretical integration. We believe it is possible that the phenomena of negative priming, inhibition of return, and cognitive inhibition more generally, may reflect different facets of a single type of underlying control mechanism (Houghton, 1993). To make this case convincingly, we need to show in detail what such a mechanism might look like and how it would function in specific circumstances to generate the attested data.

As described, our model asserts that a central function of selective attention is in the goal-based modulation of perceptual input to schemas guiding action and thought. This conception is shown diagrammatically in Figure 2.

We suppose the visual system to be delivering up multiple object-based representations of the external world in parallel. The representations contain a variety of information regarding location, form, identity, and so on. The informational content of each object representation is bound together to form a unity distinct from other entities. If some part of this information changes, it remains linked to the same object representation (or "object file"; Kahneman & Treisman, 1984), so that, in the case of a change of positional information, the entity is perceived to move rather than to be a series of distinct objects at different locations. (What transformations an object can undergo while still being perceived to be the same object, even, say, with a changed identity, is clearly an important empirical and theoretical issue, but one which we will not address here. We simply take it for granted that the visual system automatically attempts to organize spatio-temporally distributed features into objects on the basis of grouping (gestalt) principles developed through evolution.)

On the motor side, we follow other authors (e.g., Norman & Shallice, 1986; Arbib et al., 1987; Arbib, 1990) in proposing that many basic motor responses are packaged in the form of motor "schemas," such as reaching and grasping, naming, pointing at, and so on, by which the

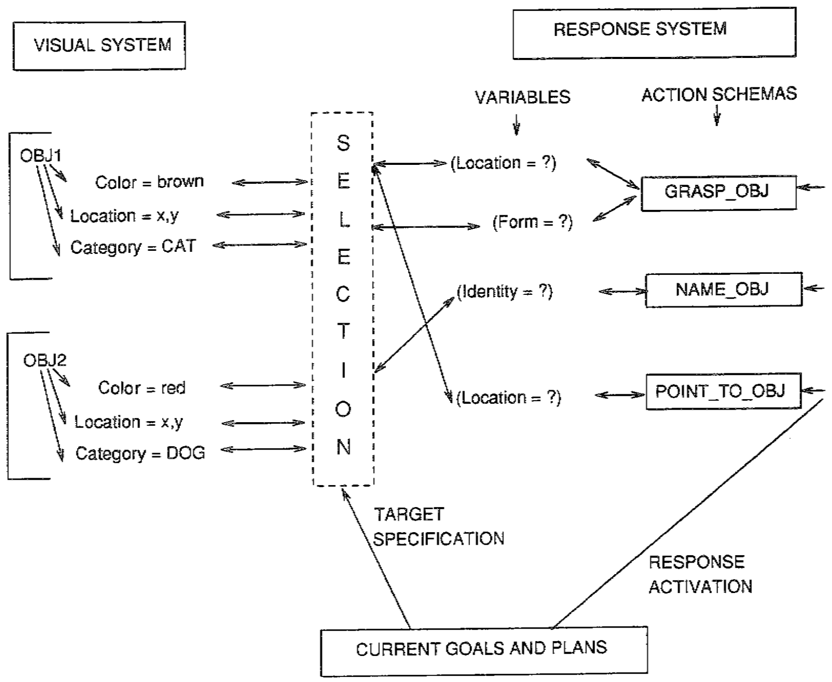

Figure 2

Schematic outline of the model showing attention as a gating mechanism at the perception–action interface. Attentional targets are specified top-down in relation to current goals and plans. These higher-level planning systems are also responsible for the activation of action schemas which have variables associated with them (Norman & Shallice, 1986). Variables are bound from information specified in perceptual representations (though other sources of binding information, e.g., memory, are possible). The attentional mechanism is object centered and foregrounds information contained in specific object representations, although selection can take place on the basis of featural targets.

agent acts on particular objects in its environment. These schemas are activated or suppressed by central planning and execution systems (Norman & Shallice, 1986). Schemas, representing general classes of actions, have parameters associated with them which, on any occasion of use, must be bound to information derived from perceptual or memory systems (Arbib, 1990). The binding of the schema parameters adapts the action to the details of its current goal and context. Thus the actions of grasping and lifting need information concerning (at least) the location (relative to the effector), form, and likely weight of a target object. We refer to this setting of variable parameters in action schemas as *binding*. We propose that information capable of binding these variables potentially will be available from a number of candidate objects in the

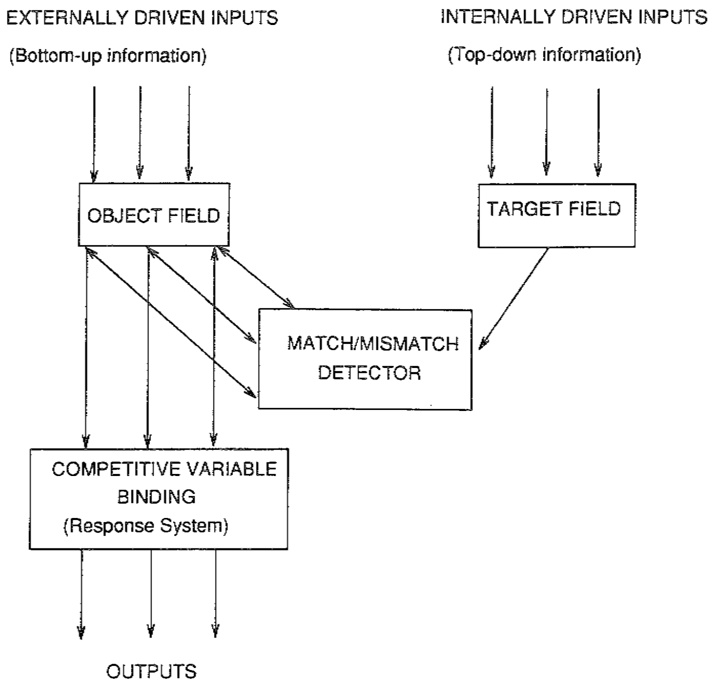

Figure 3

The interrelations between the four subsystems in the implemented model.

visual field. Coherent action toward a particular object requires that all the relevant parameters of the response schema be bound to properties of that object alone. We will refer to this *coherent* binding of a set of response variables to target objects as *coupling*. Our model proposes that successful perceptuo-motor coupling requires the intervention of a strategically controlled selective attention mechanism which sets up goal-derived targets. Information from objects matching the attentional target is selectively enhanced, whereas information from nonmatching objects is suppressed. We propose that the process of variable binding is a parallel competitive one in which all relevant information from analyzed objects feeds into the binding arena in a cascade fashion. The foregrounding of information connected to the target object provides it with sufficient advantage in the competition for binding that responses become successfully coupled to the target object.

We will now describe the selection mechanism in more detail. The principal components of the mechanism are (1) the object field, (2) the target field, and (3) the match/mismatch field. Their overall organization is shown in Figure 3.

The basic mode of action in the model is that information from both (internally driven) targets and (externally driven) perceptual object representations meet in the match/mismatch field. This system generates signals that feed back into the object field leading to the foregrounding of any objects matching the target specification. The following description of the model begins with an informal account of the organization and function of the various fields of the model, followed by a mathematical specification of the implementation used in the simulation studies.

A. Object Field

We model object representations to a first approximation as a set of linked nodes or units, as familiar from connectionist theory. Each unit represents some or other aspect of the informational content of the representation. We assume these object representations are built up automatically by the visual system in its normal functioning, the various "features" of an object being bound together to form a unified representation. The novel proposal of this model compared to typical connectionist formulations, is that the components of the object representation have an "opponent processing" type character.[2] It is commonplace in network models utilizing lateral inhibition (Grossberg, 1980; Kohonen, 1984) that individual units generate excitatory feedback onto themselves. Our model includes this self-excitatory feedback (generated via an excitatory "on-cell"), but supplements it with an inhibitory feedback loop generated by an inhibitory "off-cell." Thus, when activated, each node coding for some aspect of an object (a "property" node) sets up a combined excitatory–inhibitory feedback signal from its associated on- and off-cells. (Given the function of these cells in the model, we refer to them collectively as the *gain-control subsystem.* The basic circuit is shown in Figure 4.

We refer to the inhibitory feedback loop as the *off-channel*, and the excitatory loop as the *on-channel.* Likewise, we shall refer to the gain-control unit in the off-channel as the *off-cell*, and the analogous unit in the on-channel as the *on-cell.* In the model, we propose that these feedback signals are roughly equally weighted so that, left to itself, the gain-control feedback sums to 0.

We propose that units representing properties of the same object are linked by excitatory weights, representing the fact that the perceptual system has grouped these properties together to form the object representation. We further propose, however, that links have been formed among units in the gain-control subsystem, so that within a given object representation all the off-cells excite each other and inhibit the on-cells, and all the on-cells excite each other and inhibit the off-cells (Figure 5).

When the property cells of such an assembly are activated, they generate a stable pattern in the gain-control subsystem that echoes the pattern in the property cells, neither enhancing nor suppressing it. However, as we shall see, if some additional signal increases the activation of some subset of on-cells (even just one), this will enhance the activity of all the property cells within the object representation by the familiar process of spreading activation (and, in this case, inhibition of off-cells). On the other hand, if an on-cell is selectively inhibited (or the

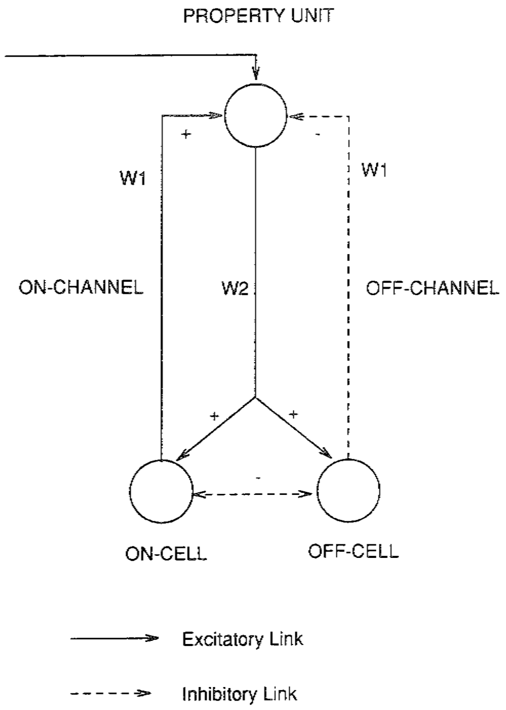

Figure 4
A single opponent-circuit. Activity in a property unit (representing the presence of some feature in the input) activates two balanced feedback circuits, one excitatory and one inhibitory (the gain-control subsystem, GCS). The activity level of the property cell can be suppressed or enhanced by changing the balance of activity in the GCS. The dependence of the activity in the GCS on the activity level of the property cell means that the feedback intensity is automatically controlled. The symbols w_1, w_2 represent model parameters.

---

[2] The general idea of opponent processing has a long history in neurophysiology and psychology, for instance, in vision (Hurvich & Jameson, 1957), emotion and motivation (Solomon & Corbit, 1974), and learning and memory (Schull, 1979; Wagner, 1981). Grossberg (1980,1983) suggests a neural net realization of an opponent-type mechanism that he terms a *gated dipole.* The dipole is differently constructed from our opponent circuit, but shares a number of functional properties, including the ability to show "rebound" behavior.

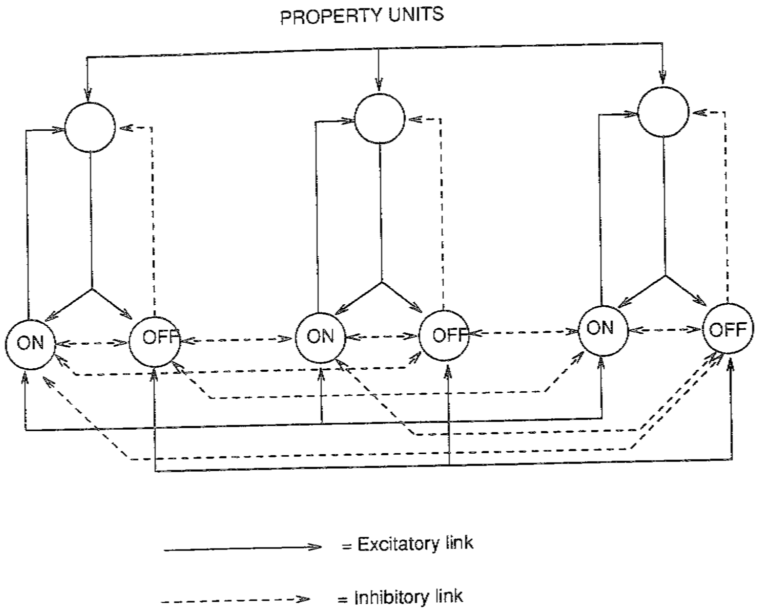

Figure 5
A set of units in an "object assembly." An object assembly is made up of a linked set of opponent circuits of the type illustrated in Figure 4. The links between the on- and off-channels of the opponent system allow for the spread of activation or inhibition through the assembly. Thus, for instance, the representation of a whole object can be suppressed if just one of its component parts is inhibited.

activation of its companion off-cell is enhanced), then, by an analogous spreading inhibition mechanism acting through the linked off-channels, the activity of all the property cells within the object representation will be reduced. As we will show, this foregrounding–backgrounding mechanism is self-stabilizing, so that when an object assembly in the "balanced" state is perturbed by an external influence to the gain-control subsystem, it will move toward a new equilibrium pattern of activation and stay there, as long as the perturbation remains. Importantly, the equilibrium pattern may be anywhere within the space allowed by the activation range, rather than just at its extremes (i.e., with all units at maximal or minimal values).

B. Target Field

Activation of nodes in the target field is taken to represent the properties of an internally generated attentional target (or template). The

content of the target specification is taken to be derived from an organism's current goals. Though able to represent sensory qualities of to-be-attended items (such as color), the target field is strictly separate from the sensory fields involved in making up the object representation. We intend that this architectural separation correspond to a physiological separation in the brain, with the target field corresponding to areas of the prefrontal lobes. A similar notion appears in the work of Näätänen (1985), who refers to what we call the target specification as the *attentional trace.*

The specification of the target leads to the biasing of the influence of perceptual information on response systems, allowing coupling of response parameters to properties of individual objects. In this way, internal goal states mediate perceptuo-motor interactions. In an experimental context, activation of targets will be due to the subject explicitly following the experimental protocol. For instance, if the subject is told to name the red object, then the activated schema in Figure 2 will be NAME-OBJ and the target property will be "red." In the implemented model, the target is thus a preset level of activity in units representing the target property (see Figure 6).

C. Match/Mismatch Field

The model proposes that the influence of top-down attentional targets on perceptuo-motor interactions is mediated by the generation of match/mismatch signals from a system which receives inputs from both internal target specifications and from perceptual systems (Figure 6). Objects having properties that fail to match the target specification (on the appropriate stimulus dimension) receive feedback to their gain-control subsystem which "breaks the symmetry" between excitation and inhibition (in the on–off channels), shifting the balance in the inhibitory direction. This inhibitory shift begins in the circuit representing the property on which selection is made (e.g., color), but by the spreading inhibition mechanism described earlier leads to the suppression of all the properties associated with the object. Objects matching the target receive an excitatory boost which begins in the circuit coding the target property and spreads via the on-channels throughout the object representation. A role for match/mismatch systems in goal-driven attention has been posited by a number of authors (e.g., Näätänen, 1982, 1985; Duncan & Humphreys, 1989; Humphreys & Müller, 1993). Näätänen (1985) reviews a number of studies of event-related potential (ERP) supporting the existence of matching operations in selective attention. Näätänen (1985) postulates that a component of the ERP known as the *processing negativity* is "generated by a cerebral matching process

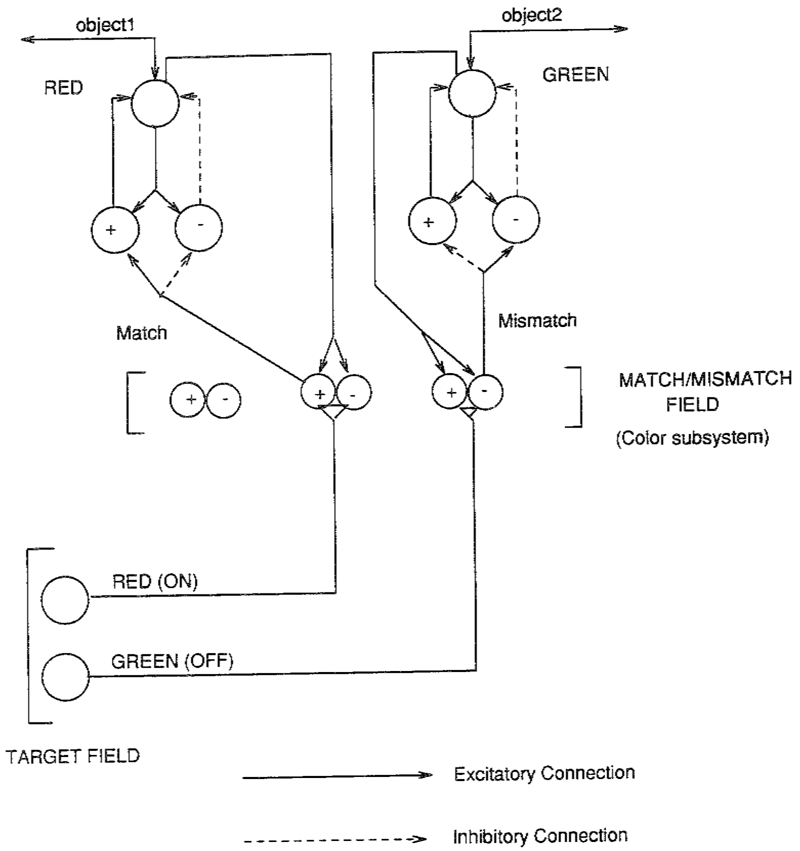

Figure 6
Figure showing the links between the object, target, and match/mismatch fields. Two objects are being represented, one red and one green (only the color circuits of the representations are shown). The attentional target is red, leading to the target unit representing red to be on. The target unit representing green is off. Signals from color cells in the object and target fields met in the match/mismatch field generating feedback signals to the gain-control systems of the object representations. Entities with properties matching the target specification have the balance in their gain-control systems shifted in the positive direction, whereas mismatching entities are shifted in the negative direction, leading to the rapid foregrounding of target objects. It should be noted that match/mismatch signals are held to be generated only from within a subfield related to the featural dimension of the target (color in this case).

between the sensory input and the attentional trace" (p. 366). In the modeling domain, Humphreys and Müller (1993) describe a network model of visual search that involves the matching of perceptual inputs with an internal template. One important reason for postulating the existence of matching operations (rather than a simpler mechanism based on direct interactions between target and perceptual systems) is that organisms "know" when they have detected targets—search behaviors are interrupted and other actions (approach, retreat, etc.) are initiated. Signals from match systems can thus play a role in regulating behavior.

The match/mismatch field (MMF) is divided up into a number of property subfields corresponding to subfields in the object property representation (for instance, a color subfield, a location subfield, a shape subfield, etc.). A target property such as *red* will be matched against specific inputs to the color subfield from the object representations (Figure 6). A match causes feedback to the gain-control system of the matching object, which leads to an excitatory imbalance in the gain-control feedback. A mismatch causes an analogous inhibitory imbalance. No signals are generated from match/mismatch subfields which code for properties of a different type than the target property (e.g., color targets are not matched against the location properties of perceptual inputs).

D. Response Binding

Inputs to the response-binding field are derived from property nodes in the object field. Recall that the idea behind the binding field is that response (action) schemas are associated with variable parameters which must be given specific values to guide the precise execution of the response. Many variables which might be bound simultaneously (e.g., form and position) are likely to be independent (uncorrelated) in an organism's experience, and must be supplied independently from the same target object. The object-based selection, which the current model postulates, allows properties to be selected in coherent bundles, although a target may be selected from distractors on only one individuating dimension (examples follow). This raises the important question of how variable binding is to be represented in a network. In the present model we use the brute force method of assuming that for any response variable there exists a set of nodes capable of representing all values of that variable. Within connectionist terms, this representation might be "distributed," with different values being represented by nonorthogonal (overlapping) patterns across the set of units, or "local," with orthogonal patterns representing each value. In the simulations presented here, the output variable is the category of the target needed by a naming response. Each category is represented orthogonally by a single node, and the model learns associations between patterns of activity over the property nodes of the object field and nodes representing (reportable) categorization decisions. Copresence of multiple objects leads to multiple activation of categories, and hence to competition within this field of category nodes. We will show how the selective foregrounding–backgrounding achieved by the model provides the object matching the target with sufficient advantage in this competition to determine the categorization response. Details of the construction of the links between the property and (response) category units are given in the discussion of the simulations.

E. A Neuropsychological Interpretation of the Model

We intend the various components of the model just described to correspond to anatomically localizable neural systems. Any such mapping from theoretical to neural systems must of course be tentative, but we believe there are reasonable grounds for suggesting the following arrangement. The object field we take to correspond to the large areas of posterior cerebral cortex responsible for the coding of visual information. This coding appears to be distributed over numerous parieto-occipital areas, some dedicated to particular stimulus dimensions (e.g., color in V4). At the highest level, regions in the inferior temporal lobes are also implicated (Mishkin, Ungerleider, and Macko, 1983). As yet unknown grouping mechanisms must act to bind these distributed codes into coherent perceptual unities (objects). We suggest that the gain-control subsystem (on- and off-cells) may be implemented locally within the cortex or may involve interaction with the thalamus by a cortico-thalamo-cortical loop regulating cortical activity. On neuroanatomical grounds, the latter proposal implicates the pulvinar nucleus of the thalamus.

We locate the target field in prefrontal cortex, which is seen by many authors as the center of nervous system "self-control," responsible for the highest levels of planning and goal-directed behavior (see, e.g., Luria, 1973; Fuster, 1980; Stuss & Benson, 1986; Goldman-Rakic, 1987; Shallice, 1988). Such ideas invariably involve a central role for selective attention (Fuster, 1980, 1989). Areas of the prefrontal lobes (e.g., principal sulcus, inferior convexivity) are known to be able to maintain visual representations (Fuster, 1980; Goldman-Rakic, 1987), and these areas have reciprocal connections with the preceding posterior visual areas (Goldman-Rakic, 1988). A number of authors have speculated that these areas are involved in the integration of perceptual information in guiding nonroutine behavioral responses, that is, those responses requiring attention (Fuster, 1980; Goldman-Rakic, 1987).

Motor schemas are generally associated with regions of the frontal cortex, such as the premotor cortex and supplementary motor area (Arbib, 1990), and we suppose that the binding of response parameters must take place in such areas. Our model thus proposes that the throughput of sensory information from posterior to frontal (motor) regions is gated by prefrontal selection targets. The remaining issue is that of precisely how prefrontal and posterior systems interact, a notable gray area in all of the preceding works referred to. Our model postulates that the interaction takes place via a matching process, and it seems possible that this could take place by direct cortico–cortical interactions between prefrontal and posterior systems, utilizing the reciprocal pathways known to exist (Goldman-Rakic, 1988). However, there appears to be other possibilities, such as an interaction with the loop from sensory cortex through the limbic system (including the hippocampus) and back to the cortex, as postulated in the visual learning model of Mishkin and his associates (Mishkin, 1982; Mishkin & Appenzeller, 1987). This model appears to attribute attentional modulation of visual learning to prefrontal interaction with this reafferent loop. This idea is consonant with our model, so long as prefrontal inputs are considered to gate activity in the loop such that activity in nommatching channels leads to suppression of associated sensory registers. In this regard it is interesting to note that Goldman-Rakic (1987) remarks that areas of the prefrontal cortex "project to the presubiculum . . . [which] represents a major output of the hippocampus to other cortical structures . . . Accordingly, the prefrontal terminals in this area are in position to gate the output of the hippocampal formation" (p. 389).

We now briefly describe the mathematical implementation of the model used in the simulation studies described in Section IV.

F. Formal Specification of the Model

The model is implemented as a neural network using components familiar from such works as Grossberg (1980), Kohonen (1984), and Rumelhart and McClelland (1986). Each unit in the model is associated with a scalar variable referred to as its *activation value.* Activation values vary in the range $[-1, 1]$ with a resting level of 0. Negative activation values represent sub-baseline states. Nodes are connected by weighted links and communicate by the spread of activation along the links. Links may have positive (excitatory) or negative (inhibitory) values. Only positive activation levels are propagated. The rule of propagation is the familiar one whereby (positive) activation values are multiplied by the weights in the connecting pathways. The *net input* to a given unit is typically the sum of the (positive and negative) inputs it receives along its input pathways. Net inputs are converted into momentary activation values by an *activation function.* The main activation function used in this model is of the "leaky integrater" type, which integrates input activations over space and time but "leaks" due to passive decay.

The function is used by all the units in the object field (property units, and on- and off-units) and by units in the response-binding field. It is given formally by

$$a_i(t+1) = \begin{cases} \delta a_i(t) + (1 - a_i(t))f(net_i(t)) & \text{if } net_i > 0 \\ \delta a_i(t) + (1 + a_i(t))f(net_i(t)) & \text{if } net_i < 0 \end{cases} \tag{1}$$

where a_i is the activation level of unit u_i, and net_i is the net input to u_i, and δ is a "decay" parameter. Because we use negative activation values to represent states of suppressed responsiveness (which should not be thought of as the inverse of excited states, or as the possession of "negative activation"), we in fact use two "decay" parameters: δ^+, representing passive decay from excited states (i.e., $\delta = \delta^+$ if $a_i > 0$), and δ^-, representing spontaneous recovery from suppressed states (i.e., $\delta = \delta^-$ if $a_i < 0$). In *all* the reported simulations $\delta^+ = 0.5$ and $\delta^- = 0.9$. To ensure that spontaneous recovery being slower than passive decay. To ensure that activations remain bounded, the function $f(x)$ should be some suitable *S*-shaped "squashing" function, with output in the range $[-1,1]$, and the property $f(0) = 0$. In the reported simulations we use

$$f(x) = \frac{2}{1 + e^{-x}} - 1 \tag{2}$$

which has an appropriate sigmoidal shape with asymptotes at -1 and 1. We note in passing that we have run the model using other activation functions of a similar type, and have not found it to make any difference to the phenomena discussed here.

Because negative activation values represent states of suppressed responsiveness, such values clearly are not propagated. Thus each unit has an output function,

$$o_i(t) = \max(0, a_i(t)) \tag{3}$$

that is, the output is 0 unless the activation value is positive. This will be assumed in all the following equations, where a_i should be taken to represent the output activation value [i.e., $a_i(t) = o_i(t)$].

The behavior of different nodes depends essentially on their net inputs from outside sources. In the assemblies of nodes forming object representations, we have three types of node (see Figure 4): u^p, representing some property, u^{on}, the on-cell, and u^{off}, the off-cell. Their net inputs, net^p, net^{on}, and net^{off} are given respectively by

$$net_i^p = I_i^{ext} + w_1(a_i^{on} - a_i^{off}) \tag{4}$$

$$net_i^{on} = w_2 a_i^p + w_3 \sum_{j \in A} a_j^{on} - w_4 \sum_{j \in A} a_j^{off} + mm_i^{on} \tag{5}$$

$$net_i^{off} = w_2 a_i^p + w_3 \sum_{j \in A} a_j^{off} - w_4 \sum_{j \in A} a_j^{on} + mm_i^{off} \tag{6}$$

where I_i^{ext} is the external input to property unit u_i^p; a^p, a^{on}, and a^{off} are the activations of the property, on- and off-units respectively; w_1, w_2, w_3, and w_4 are weight parameters; A is the set of nodes in the assembly; and mm_i^{on} and mm_i^{off} are signals from the match field [defined in Eqs. (11), (12)]. The external input signal I_i^{ext} is derived from sources outside the model and is set to a value of 1.5. Given the activation function used [Eq. (1)], this is sufficient in itself to drive the activation of a property node to a value of about 0.75. Therefore, in the simulations reported in the following section, the establishment of property node activation levels above or below this value is due to the selection mechanism.

The parameters w_1 and w_2 govern the strength of the interaction between property units and the gain-control cells, and vice versa (Figure 4). Thus, w_1 is the magnitude of the weight of the feedback links from the on- and off-cells to the property unit in an opponent circuit, and w_2 is the weight from property units to gain-control units. Having just one weight parameter in each direction (instead of two independent ones) means that the circuit is inherently "symmetrical" (though this is, in fact, not strictly necessary for the model to function). In all the simulations, $w_1 = 1.3$ (except where it is purposefully manipulated) and $w_2 = 1$. We have not systematically investigated the effects of changing the relative values of these weights or of using nonsymmetrical weights.

The parameters w_3 and w_4 govern the strength of the lateral interactions among the gain-control (on–off) cells. Their values are not fixed but depend on the number of property nodes in the object assembly. These weights control the strength of the feedback within the on–off channels, and it is important that the total feedback any node receives via these channels does not grow too large. Otherwise, for instance, the activity in a channel could become self-sustaining because of the positive feedback. It is also desirable in the implementation to keep the *total* input weight to any gain-control cell from other cells within an assembly constant, independent of the number of other such cells that happen to be in the assembly. This means that the strength of the feedback any gain-control cell receives is independent of the number of cells in the object assembly. Normalization can be achieved locally if each cell is thought of as having a total input "weight," which it distributes evenly among its inputs. Thus the actual values of the parameters w_2 and w_3 are given by

$$w_3 = \frac{W_{tot}^+}{|A|} \tag{7}$$

$$w_3 = \frac{W_{tot}^-}{|A|} \tag{8}$$

where $|A|$ is the number of nodes in the assembly. The parameter W_{tot}^-

represents the total inhibitory weight to a cell (i.e., from off-cells to on-cells, and vice versa), and W_{tot}^{+} is the total excitatory weight (from on-cells to on-cells, and off-cells to off-cells).

Units in the match/mismatch field receive input from object property units and from target units. The field contains two classes of units: match units and mismatch units. Only units coding properties of the same class as the current target property can be activated; for example, if targets are specified by location, only the location subfield of the match/mismatch system will be active. Match units of the active subfield fire when they receive inputs from both object and target field cells (i.e., target and property are the same). Mismatch units fire when they receive inputs only from property cells (target and property are different). In the implementation, the activation values of match/mismatch cells are proportional to the strength of the input from the property node, modulated by the strength of the target activation. Formally, for all match/mismatch units u_i^{m+} and u_i^{m-} in the appropriate selction subfield,

$$a_i^{m+} = a_i^{tar}\, a_i^{p} \tag{9}$$

$$a_i^{m-} = (1 - \mathrm{sign}(a_i^{tar}))a_i^{p} \tag{10}$$

where a_i^{m+}, a_i^{m-}, a_i^{tar}, and a_i^{p} are the activations of the *i*th match, mismatch, target, and property units, respectively (i.e., for simplicity, units at corresponding locations in the property, target, and match fields code for the same feature). From Eqs. (9) and (10), we see that when $a_i^{tar} = 1$, then $a_i^{m+} = a_i^{p}$ and $a_i^{m-} = 0$. That is, when a target node is active, the strength of the match signal (a_i^{m+}) generated from the corresponding input property node is equal to the activation level of that node. The mismatch (a_i^{m-}) signal is 0. When $a_i^{tar} = 0$, then $a_i^{m+} = 0$ and $a_i^{m-} = a_i^{p}$. That is, when there is no target input, the mismatch signal is equal in magnitude to the activation of the property node, and the match signal is 0. When $a_i^{p} = 0$ (i.e., there is no sensory input to a match/mismatch channel) no signals are generated, whether there is target node input or not, that is, sole activation of a target specification does not give rise to activation in corresponding sensory registers (property nodes). This assumption is probably incorrect, judging by physiological indicators of heightened activity in cortical regions corresponding to a to-be-attended stimulus, even in the absence of the stimulus itself (see, e.g., Roland, 1982). The model could be altered to reflect this "expectancy" priming (which would make selection more efficient), but this would not appear to have any bearing on the issues which concern us in this chapter.

The match/mismatch signals project to the gain-control cells associated with the property unit whose activation caused the match or mismatch. Match signals inhibit the off-cell and excite the on-cell, mismatch signals inhibit the on-cell and excite the off-cell. Formally,

$$mm_i^{on} = a_i^{m+} - a_i^{m-} \tag{11}$$

$$mm_i^{off} = a_i^{m-} - a_i^{m+} \tag{12}$$

where mm_i^{on} and mm_i^{off} are the match/mismatch inputs to the i^{th} on- and off-cells, respectively [see Eqs. (5), (6)] and a^{m+} and a^{m-} denote the activation levels of match and mismatch cells, respectively [Eqs. (9), (10)].

The following tabulation lists the given model parameters and provides typical values used in the following simulations:

$I_i^{ext} = 1.5$

$\delta^{+} = 0.5$

$\delta^{-} = 0.9$

$w_1 = 1.3$

$w_2 = 1.0$

$W_{tot}^{+} = 1.3$

$W_{tot}^{-} = 1.3$

IV. MODEL SIMULATIONS

This section provides detailed examples of the dynamical properties of the model and simulates data from a number of paradigms, including interference effects and negative priming.

A. Selection and Interference

The simulations in this section illustrate the basic dynamics of selection in the model. The case considered is the generation of an overt categorization response (naming) to one of two simultaneously presented line drawings of common objects. The naming target is distinguished from the distractor by its possession of some criterial attribute—this attribute will generally be the object's color. (This is the display type used in, e.g., Tipper, 1985.)

To implement such simulations, the network was trained to associate sets of properties with object categories. In terms of Figure 3, weights were learned between distributed (i.e., correlated) patterns of activity in the object field and activity in single nodes representing categories in the response system (Figure 7). These weights were learned using the

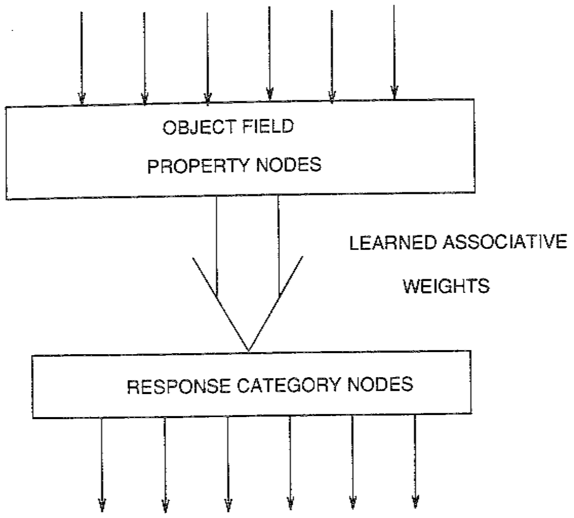

Figure 7
Weights linking nodes in the object and category fields are learned using the delta rule supervised learning algorithm. After learning, a distributed pattern of activity in the object field, representing a single object, gives rise to a single (orthogonal) categorization decision. The presence of multiple, simultaneously analyzed objects can give rise to competing bindings of the categorization response.

delta rule associative learning procedure (Rumelhart & McClelland, 1986), though the model is not tied to the use of this particular algorithm. After learning, presentation of an object is represented by activity over a particular set of property nodes. This activity feeds through the learned weights generating a pattern of activity on the category nodes, considered here to represent the binding of the naming response to a particular category (Figures 2 and 3). If multiple objects are present in the input, it will be expected that multiple category bindings will be coactivated and that selection, therefore, will have to take place to generate a coherent response to the target.

The property nodes are organized into a number of property subfields, according to stimulus dimension. The dimensions used include location, position, size, shape, animacy, and function. These dimensions are important in the matching process, as match/mismatch signals are only directly generated within the selection dimension subfield. Some of the nodes in the object field stand for contingent properties of the particular object presented (e.g., color and position). These properties, although bound into the object assembly, are irrelevant with regard to the categorization response. Other activated nodes stand for defining characteristics of the object (i.e., characteristics present during training of the property to category mapping). Patterns of activity over such property nodes are assumed to encode semantic (categorization-relevant) features, and may be thought of as a distributed representation of the semantic classification of the object. In the following simulations, object assemblies consist of nine or ten property nodes, of which five are always category relevant, the remainder code for color and position. The generation of an overt (conscious) classification response thus involves the mapping of the distributed semantic representation (in which different objects of correlated categories might share nodes) onto a local representation of category (in which different categories are orthogonal to each other). Thus nodes representing, say, *ANIMATE HUMAN MALE TWO-LEGS*, would give rise to a peak of activity in a category node representing the category *MAN*, with relatively little activation of the other category nodes. In the simulations reported here, there were 50 property nodes in the object field and 12 category nodes in the response (parameter-binding) field. Given the replication of the property node structure in the gain-control and match fields, this yields a total of 262 nodes.

We now illustrate the way in which the mechanism just described realizes the central function of selection of objects and the concomitant coherent binding of response parameters in situations in which a target object is accompanied by one or more distractor. In the following simulations the objects presented will be thought of as colored line drawings, with the selection cue being color (Tipper, 1985). The simulations are produced by activating property nodes for both objects (say, a red cat and a green guitar), which are linked in two assemblies, as previously discussed. The process of linkage (object grouping) is not explicitly modeled here, but is assumed to take place during lower level perceptual processes (as stated earlier, this model of attention is object based). Concurrently, in the target field, a node representing the target property (say, the color red) is activated. Activation from the object field feeds through the learned connections to the category field (as discussed previously) giving rise to two competing response tendencies. Selection is achieved by virtue of the fact that internal targets and externally driven activations come together in the match/mismatch field. In the particular example under consideration, the target *red* generates a match with the property "red," which forms part of the object assembly representing the cat, and a mismatch with the property "green," which forms part of the guitar representation. As discussed previously, the match/mismatch signals break the symmetry in the gain-control feedback within the object field. The match signal from *red* causes the opponent mechanism to generate net excitatory feedback to the property "red," which spreads through the relevant object assembly (via the on–off nodes), giving a boost to all property nodes within it. Conversely, the mismatch with the

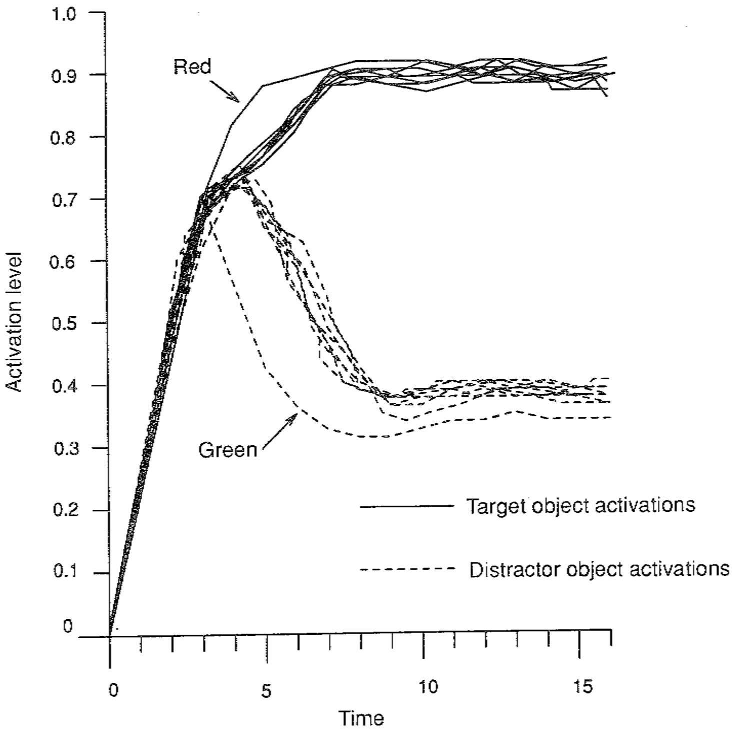

Figure 8
Attentional separation of target and distractor objects using color as the dimension of selection. As input arrives in the object field the two assemblies briefly attain similar activation levels, but eventually the nodes representing target (RED) and distractor (GREEN) colors are respectively enhanced and suppressed by virtue of matching or failing to match the internal target (RED). This activation gap then spreads through to the rest of the property nodes in the respective assemblies by the linked on–off channels described in the text.

property node "green" tips the balance in its gain-control feedback in favor of the off-channel. This heightened off-activation spreads from the node coding for green to the rest of the relevant assembly via the lateral connections between the gain-control units (Figure 5).

To illustrate this process, Figure 8 shows the activation values of nodes in the object field when two objects are simultaneously presented, one of which is to be responded to (selected) on the basis of possessing some criterial property (color, in this case). The target object is in the color *red* and the distractor is in *green.* At stimulus onset, all activated property nodes increase their activation levels together. At a certain point, however, the feedback from the match system starts to

influence the pattern of activity. First of all, the nodes which give rise to the match/mismatch signals (*red* and *green*) become separated, *red* being amplified and *green* suppressed. This amplification/suppression rapidly spreads to the rest of their respective assemblies, leading to the separation of the two sets of nodes.

It should be noted at this point how the model does not cause the representation of the distractor to fall below background levels (i.e., to disappear). This is partly due to the fact that the activation of the property nodes drives the activity in the gain-control feedback channels—as the activity level of a property node falls it generates less inhibitory feedback onto itself. The resting level represents the equilibrium activation between the external driving input and the self-generated inhibitory feedback. The model thus provides an implementation of an automatically self-adjusting gain-control mechanism, which allows the strength of the inhibitory feedback to be continually responsive to distractor activation levels.

The effect of this separation of target and distractor inputs on responding is shown in Figure 9. To begin with, there is an equal tendency to bind the response to both input patterns, preventing any response being made (two categories cannot be named at once). However, as the target and distractor activations in the object field become separated, this leads to one categorization decision becoming dominant. Although we have not attempted to implement a specific response generation mechanism, we suppose that there is some minimal dominance of one binding over another, which is required before a response can be confidently initiated. Thus, reaction time will correlate positively with the time taken for a clearly dominant response binding to emerge. These simulations show how the model implements the idea that the selective attention acts to bias the competition for control of the parameters of the response systems.

The strength of the gain-control feedback in the on–off channels, which leads to these results, is dependent not only on the strength of the input from the property nodes but also on the values of the weights in the on- and off-channels, which are set to be equal (in magnitude) and opposite [parameter w_1 in Eq. (4)]. From this it seems clear that the degree of separation achieved at asymptote between target and distractor will be a function of the magnitude of these weights. Because it is this activation advantage which allows a dominant response tendency to emerge, it follows that this selection gradient will be reflected in the response system.

This is illustrated in Figure 10, which shows the timecourse and asymptotic values of the activation difference between target and distractor for different values of the magnitude on the feedback parameter w_1. As can be seen, the greater the value of w_1, the greater and more

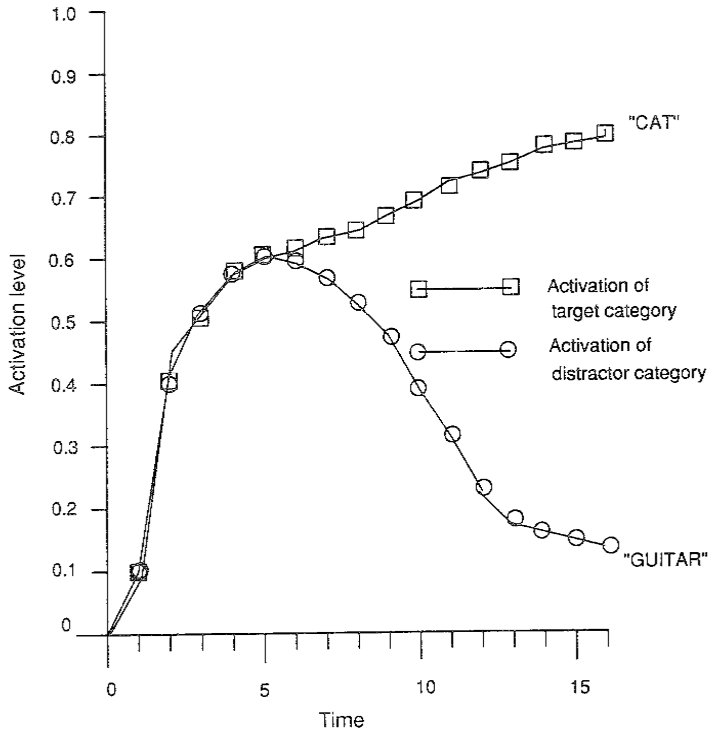

Figure 9
Activation levels of nodes representing the binding of the identification (naming) response to a particular category in response to the object field activation pattern shown in Figure 8. The attentional separation of target and distractor representations leads to the dominance of information from the target in the categorization decision. Because this information is "consistent," a clear categorization emerges.

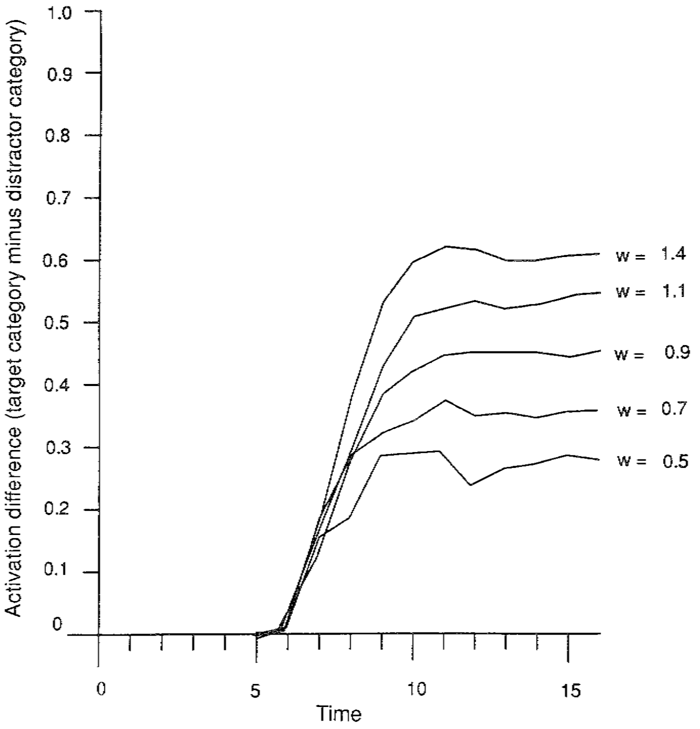

Figure 10
Difference in activation level between target and distractor categories (response binding) over time for different magnitudes of the opponent-circuit feedback parameter w_1. Higher values of w_1 produce more rapid separation of target and distractor categories as well as higher asymptotic values.

rapid is the separation achieved between response strength to target and distractor. In other words, higher values of w_1 lead to more efficient selection. As previously discussed, there is known to be wide individual variation in selection ability, as well as systematic differences between various groups. The model thus suggests possible mechanistic sources for these differences. There are other manipulations which may be made, giving a variety of individual selection profiles. For instance, for the model to function, it is not necessary that the feedback weights (w_1) have equal magnitudes in the on–off channels. Greater strength in the off-channel, for instance, leads to an overall dampening effect.

A familiar finding in studies of interference is that greater similarity beween target and distractor leads to greater interference (Estes, 1972; Bjork & Murray, 1977; La Heij, 1988; Duncan & Humphreys, 1989; LaBerge & Brown, 1989).[3] This increase in interference is found in the model if it is assumed that two separate objects can share property nodes, that is, if the same property node can be bound into more than one object assembly. Thus, in a display showing a cat and a dog, individual nodes representing properties such as *ANIMATE HAS-FOUR-LEGS* would form part of two assemblies simultaneously. These nodes would thus receive positive (excitatory) feedback from the other nodes

[3] We note that distractor interference effects are complex. Although distractor interference can be greater when the distractor is related to the target (Klein, 1964; Stroop, 1935) and such interference is determined both by semantic properties and response set (La Heij, 1988; La Heij, Dirkx, & Kramer, 1990), other work has shown facilitation from related distractors (La Heij, Van Der Heijden, & Schreuder, 1985; Underwood, 1980). One explanation is that there are multiple representations associated with the ignored object, for example, for its perceptual, semantic, and response properties. The effects of these representations may be to facilitate or inhibit performance, depending on task demands.

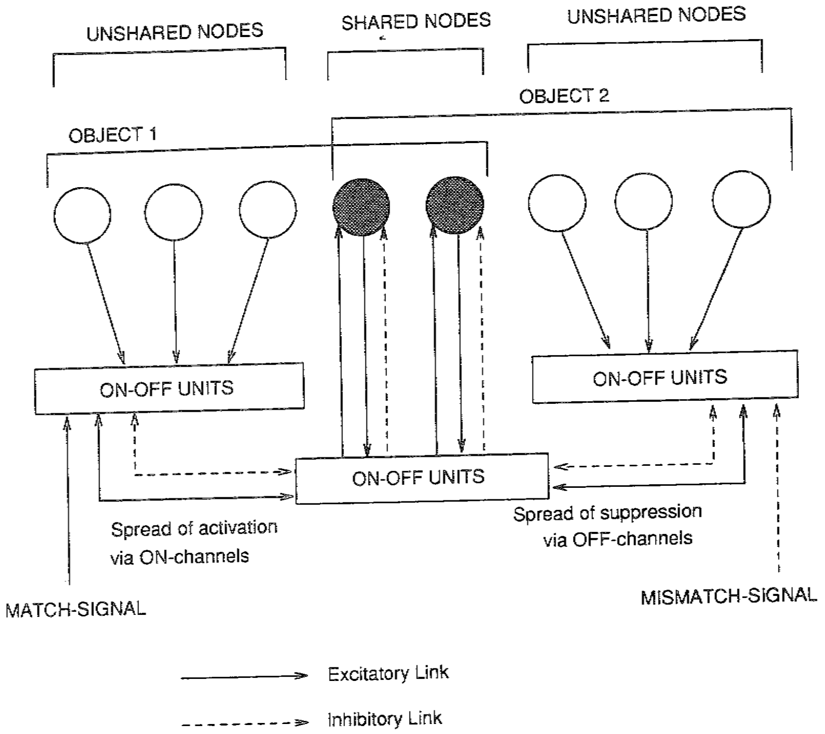

Figure 11
An example of two object assemblies sharing nodes. Each assembly consists of five nodes with two nodes being shared with the other. The assembly labeled OBJECT 1 is shown as matching a selection target (on some dimension) and thereby receiving a match signal to its gain-control subsystem. The nonmatching object (OBJECT 2) receives a mismatch input, flipping its gain-control system into inhibitory mode. The two shared nodes, being part of both assemblies, are thus subject to both amplificatory and suppressive influences spreading from the nodes that are unique to each object.

forming part of the object assembly matching the target, and negative feedback from the nonmatching object. (See Figure 11.)

This combination would be expected to yield an activation level in the shared nodes that would be higher than that achieved by the unshared nodes in the distractor, yet lower than that achieved by the unshared nodes in the target. Because the activation of these shared nodes can spread via the on-channels to other components of the distractor assembly, we predict that the general level of activation of a distractor will be a positive function of its degree of correlation with the target, thus giving rise to greater interference in the binding process.

To illustrate this phenomenon, the selection process just described was run with a series of distractor objects varying in their degree of similarity to the target and implemented as the sharing of category-relevant property nodes. As before, in each trial two object assemblies were simultaneously activated representing two objects differing in color. Selection was based on the activation of a node in the target field specifying only the color of the object to be named. All parameters were unchanged. In the simulations, the shared features between target and distractor were always of the criterial kind (i.e., the shared features were relevant to the categorization decision). Although sharing noncriterial features would be expected also to produce interference (by the same process as that described), these features themselves would not directly contribute to the activation of competing responses. Thus the more difficult case for the model is to generate a coherent response in the presence of distractors with shared criterial features, and it is crucial to demonstrate that it is capable of such performance. The dependent variables looked at were the difference in the activation levels achieved by target and distractor objects in the object field, and the subsequent level of interference in the binding process. Figure 12 shows the difference in the mean activation levels achieved by object field nodes for four levels of correlation between object assemblies. Each curve is labeled with the percentage of shared active nodes in the two assemblies, given for both criterial (category-relevant) features ($c1$) and total (criterial and noncriterial features ($c2$).

As the correlation rises the mean activation gap achieved between target and distractor objects decreases. This decreased activation gap between sets of property nodes is reflected in the category-binding field in the form of a poorer signal-to-noise ratio for the target category. This greater interference simply reflects the fact that, in this model, overt categorization is implemented as a process of orthogonalization (correlated object-field pattern vectors are mapped to orthogonal categorization vectors). The greater the correlation between two distinct simultaneously presented inputs, the more difficult this becomes (i.e., the patterns become harder to separate).

Earlier we presented two general arguments regarding why inhibition is likely to be an important component of any biological selection process. To make this point further, we emphasize that it would be impossible to achieve this separation of patterns with shared nodes without the ability to specifically suppress the features unique to the distractor (i.e., without the inhibitory component of the gain-control mechanism). If the model contained only a facilitatory mechanism boosting target activation, then enhancement of the shared properties in the target assembly would spread unopposed through *both* assemblies, leading to enhancement of the distractor and the inability to selectively respond. In general, it is difficult to see how any model which uses correlated, distributed representations (thereby allowing sharing of nodes in object

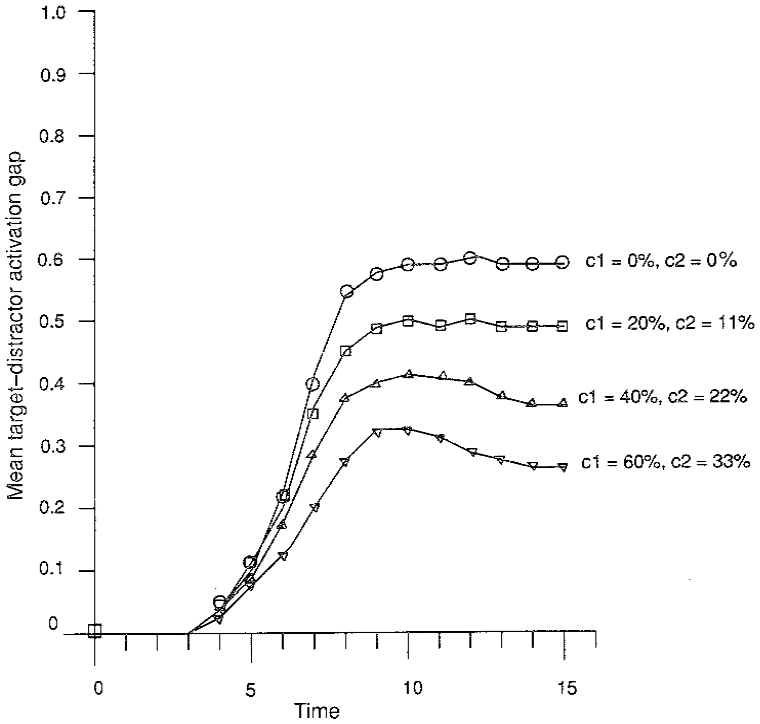

Figure 12
Simulation of similarity effects on efficiency of selection. The greater the similarity between two (categorically distinct) objects (in terms of shared nodes in the two assemblies), the greater the interference caused by the distractor, as measured here by the difference in the mean activation values achieved by the nodes composing the target and distractor object assemblies. Each curve in the figure is labeled by two percentage correlation values, *c*1 and *c*2. The former gives the percentage of category-relevant active nodes that the target and distractor share, and the latter the percentage of overall active nodes (including, e.g., nodes representing location).

representations) could achieve effective selection without the active inhibition of nontarget related features counteracting the spread of activation from the attended object. We additionally note that these features may be arbitrarily many, and, as is the case in these simulations, may not be related in any way to the endogenously generated selection feature (and thus cannot be directly inhibited by a mismatch signal). It is thus hard to escape the conclusion that the specific inhibition of such features must be due to a spreading inhibition mechanism (as proposed by Tipper, 1985), with its source in the dimension on which selection is being made.

Previous discussions of this idea (e.g., Neumann & DeSchepper, 1991) have considered spreading inhibition as analogous to spreading activation, and have not addressed the issue of how it is that inhibition *can* spread. Activation can spread because excitatory links are, so to speak, transitive. If a node u_1 excites a node u_2, and u_2 excites u_3, then activation of u_1 will eventually spread to u_3. If the links in this chain are inhibitory, however, then activation of u_1 will inhibit u_2, the effect of which will be to *disinhibit* u_3, the exact opposite of what is required. Hence, spreading inhibition cannot operate in a manner strictly analogous to the typical mechanism of spreading activation. The model presented here contains an effective solution to this problem, the only one we are aware of.

B. Negative Priming

We now consider how the model accounts for the basic facts of the negative priming phenomenon discussed earlier (Tipper, 1985). In this paradigm, a priming selection trial, which might be represented in the model by Figures 8 and 9, is followed by a probe trial. In the probe selection trial, the target item might be the same as in the priming trial (repeated prime condition), the same as the distractor item in the priming trial (ignored prime condition), or different from both (novel probe condition). It is commonly found that reaction times to the ignored prime condition (i.e., when the current target item was the previously ignored distractor) are slower than in the novel probe condition. Reaction times are fastest to the repeated prime. In the following simulations we will only be interested in the comparison between the ignored prime (IP) and the novel probe (NP) conditions. We assume that the repetition priming effect is due to factors which we do not attempt to model here, such as increased connection strength in selected perception–action pathways, or in the lowering of selected response thresholds. Repetition priming appears to last for some time (at least a matter of hours, possibly much longer) and can withstand any number of intervening events, whereas the negative priming effect does decline over time (Neill & Valdes, 1992), and appears not to survive the making of intervening responses (Neill & Westberry, 1987; Tipper et al., 1991).

Previous informal accounts of the negative priming phenomenon (Tipper, 1985; Tipper & Cranston, 1985) stipulate that during selection, distractors are actively inhibited, and the inhibition is evident in the retarded responding in the IP condition. This informal account does not stipulate *to what extent* distractor items are inhibited while they are present. For instance, are their internal representations suppressed below resting levels? Our formal model makes specific proposals in this regard. As illustrated in Figures 8 and 9, which show how the model achieves selection and coherent responding, it is not necessary for the model to "obliterate" any internal trace of the distractor object in order to respond properly to the target; that is, it is not necessary to suppress its representation below resting levels. Indeed, we do not believe it is appropriate to postulate that the distractor is suppressed below resting levels for a variety of reasons. As discussed earlier, nonattended objects in the world form a background, the representation of which facilitates both recognition of particular objects (by providing contextual information) and effective action toward them (by allowing the adaptation of actions to unpredictable details of the physical context). Further evidence that the unattended field is not only not deleted but is actively monitored comes from the orienting response. Organisms exhibit orienting responses to salient or potentially significant events in the unattended field to the extent that ongoing, useful activity (e.g., eating) may be disrupted. A particularly relevant example of orienting to the background is when we notice the *offset* of an unattended background noise, whose presence we had ceased to be conscious of. It is difficult to see how the offset of an unregistered signal is likely to be noticeable. There is also experimental evidence of distractor objects (associated with specific responses) being processed even to the point of incipient response generation, as indicated by electromyogram recordings in appropriate effectors (Erikson, Coles, Morris, & O'Hara, 1985). All such considerations support the notion that the unattended field is analyzed and continually monitored as a background to focal actions. This would not be possible if selectively attending to one part of the afferent field effectively obliterated all else. Our model thus does not postulate that the representation of distractors is suppressed below resting levels (Figure 9), and we consider such a proposal to be, in the general case, untenable.

This apparently raises something of a quandary. Suppose that Figure 9 represents the priming trial in a negative priming experiment and that at the end of the trial the display is terminated and the distractor disappears. The object field nodes constituting its internal representation now no longer receive external input. We might expect that their activation values would then decay back to resting levels, given the decay term in the activation function [Eq. (1)]. On the basis of this it would be difficult to predict any kind of negative priming when the probe display appears in the IP condition. We might even predict a modicum of facilitatory priming based on any residual activation in the relevant units at the onset of the probe display.

However, this is not in fact what happens. As noted earlier, the equilibrium activation level achieved by any property unit constitutes a balance between its net input and the spontaneous tendency to decay to

resting levels. This level is lower for units in distractor representations than for the units in targets because, in the distractor assembly, the activity in the linked opponent circuit is concentrated in the off-channel, counteracting the effect of the excitatory external input, and reducing the overall net input. Once the external excitatory input is terminated, however, the net input to a property unit in a distractor representation consists of just inhibitory feedback from the off-channel, driven by the activation of the property units themselves. We can see this from the net input equation for a property node u_i^p, Eq. (4), repeated here.

$$net_i^p = I_i^{ext} + w_1(a_i^{on} - a_i^{off})$$

Thus for a property node u_i^p in a distractor assembly, when the external input is terminated, then $I_i^{ext} = 0$, $a_i^{on} < 0$ (on-channel is suppressed during selection), and $a_i^{off} > 0$ (off-channel is boosted). Thus at the offset of the prime display, $net_i^p = -w_1 a_i^{off} < 0$. The effect of this switch of the net input from excitatory to inhibitory is to cause an inhibitory "rebound" in the activation values of the property units constituting the distractor representation. This process is illustrated in Figure 13, which shows what happens to the property nodes in the object field for a display of two objects, one of which is selected for responding and one of which is ignored.

The two curves in the figure represent the mean activation values of all the nodes in each of the two object assemblies (the assemblies are orthogonal in this case). As can be clearly seen, at the offset of the display, while the representation of the target object gradually decays back toward resting levels,[4] the representation of the distractor returns to a resting equilibrium value via a different route involving an excursion into the negative activation region representing suppressed responsiveness.[5] We propose that it is this postoffset inhibitory rebound in the

---

[4] To try to model what happens when displays are terminated, it is necessary to add two further processing assumptions to the model. The first is that the assembly of nodes representing an object percept should immediately start to decay as an assembly; that is, the lateral connections linking the on–off cells in each assembly should decay back to 0, so that the property cells are no longer grouped in an object percept. In the implementation, after stimulus offset $w_{ij}(t + 1) = 0.9w_{ij}(t)$, where w_{ij} is the weight linking any two on- and/or off-cells. If the weights do not decay, activity within, say, the off-circuit of the distractor, can be self-maintaining, generating long-term negative priming, akin to an object-shaped "hole" in perception. There would also be disastrous interference between successively presented objects which share nodes, producing a summed assembly containing all input properties. The second assumption is that at stimulus offset, activity of the target node passively decays at the usual rate. Without this assumption, decay of the target–object representation is considerably slower, due to continued match field input.

[5] This "rebound" behavior at offset of external input is typical of opponent mechanisms, and is largely responsible for their attraction in explaining a variety of phenomena (see, e.g., Solomon & Corbit, 1974; Houghton, in press).

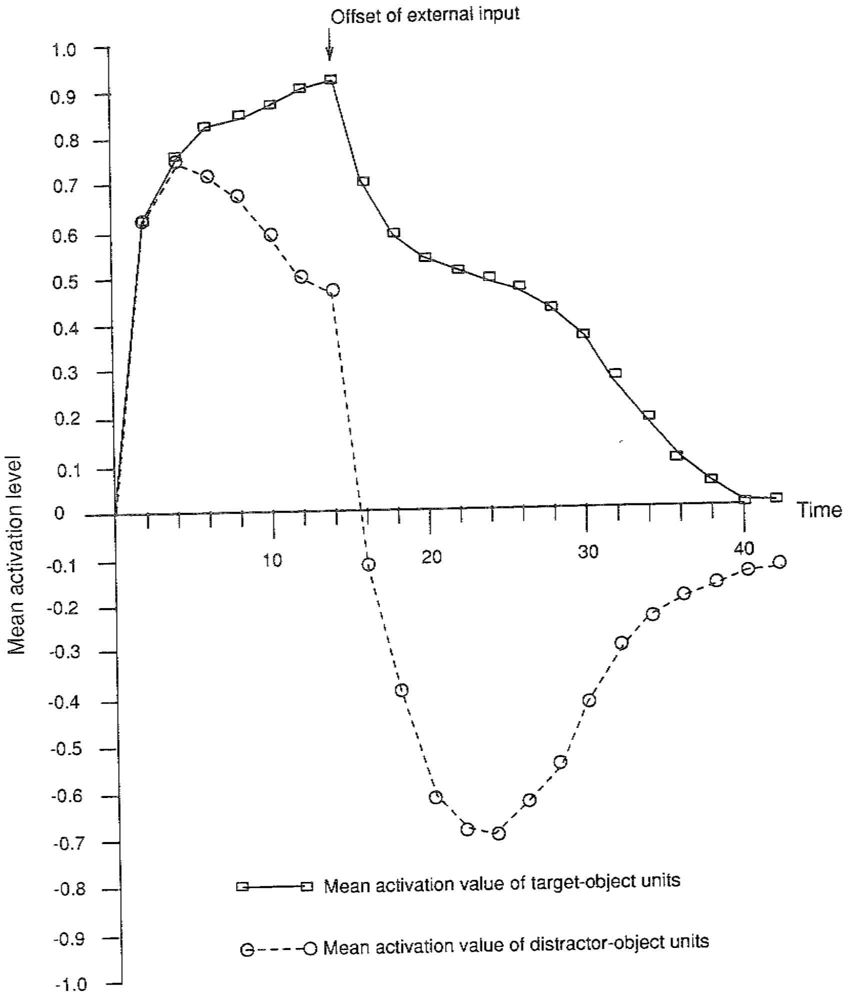

Figure 13
Showing how the model can accommodate the negative priming effect with the need for the distractor representation not to be suppressed below resting levels while there is external input. The two curves show the mean activation values over time of the units in two object assemblies, activated by an external display in a selection task. In particular, the dashed curve shows the mean activation of a distractor stimulus. At offset of the external input, the representation of the distractor does not passively decay to resting level but suffers an inhibitory rebound.

activation of the distractor representation that causes negative priming. If the distractor is re-presented as a target during the time the nodes constituting its internal representation are in this suppressed state, then

reaction times will be retarded relative to a novel probe (whose constituent nodes will be at resting level activations). Thus the model successfully accommodates negative priming with the need for the representation of distractor inputs to be maintained above resting level. Moreover, no special mechanism is required to achieve this. The inhibitory component of the selection mechanism, being based on self-feedback via the opponent system, is automatically self-regulating so that ignored distractors reach equilibrium activation levels above resting level (but below that of the target). It is this same opponent mechanism that is responsible for the postoffset rebound (negative priming) found when the distractor input is suddenly terminated (or, conversely, if attention is suddenly switched from a target; see the simulation of inhibition of return in the following section).

The effect of the inhibitory rebound at prime offset on subsequent representation in the probe is shown in Figures 14 and 15. These figures show the mean activation values achieved by target and distractor assembly nodes in a priming trial followed by a probe trial (the curves are actually the means of 10 runs of the model using different sets of input objects and some random variation of the parameters, w_1, δ^+, and δ^-). As before, each assembly contains one node representing color (red or green). The target object is red in both the prime and probe trials, and the distractor is green. One consequence of this is that in the ignored prime condition, the ignored distractor must change color. Figure 14 shows the novel probe (NP) condition in which the target in the probe trial is unrelated to either of the objects in the priming trial. In this case, there is little difference between the prime and probe conditions. The target is separated from the distractor while both are on (presentations here are fairly brief and maximal separation is not achieved while the input is still on). After stimulus offset (when responses are actually made in such experiments), there is clear dominance of the target over the distractor (which is, of course, inhibited). Response binding in the probe trial is thus not significantly different than in the prime.

By contrast, Figure 15 shows the situation in the critical ignored prime condition. Here the target in the probe display is the same category as the previously ignored distractor. In this case, the inhibitory rebound at the offset of the prime interferes with the re-representation of the stimulus as a target. Indeed, at probe onset, the distractor achieves a higher initial activation than the target which takes time to establish dominance. Interestingly, at probe offset, the decay of the target representation is more rapid than in the novel probe condition. This is because, in the ignored prime condition, the establishment of recirculating activation in the on-channels of the target is badly interfered with by the residual activation in the off-channel, established when the probe target was the previous distractor. In the novel probe condition, it is this

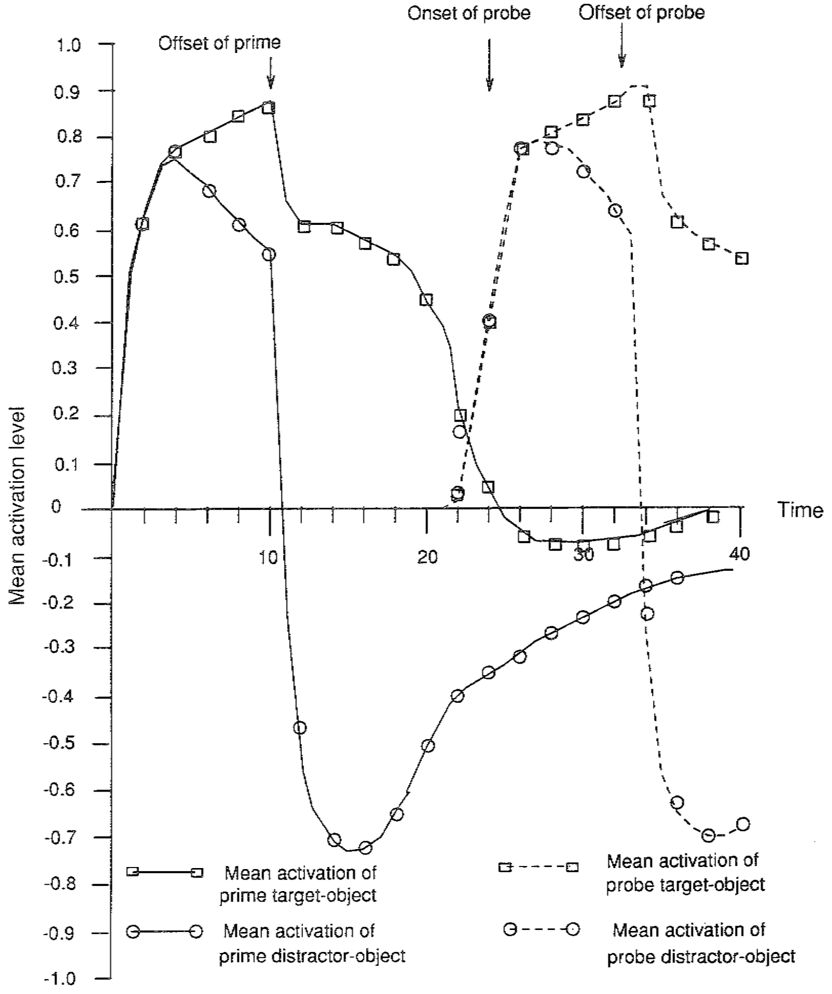

Figure 14

Full simulation of a negative priming trial with both a prime and a probe display. This figure shows the mean activation of object assembly nodes in the NP (novel probe) condition. Selection is achieved equally well in both the prime and probe sections of the task.

self-reinforcing on-channel activity that enables the target to maintain its activation after display offset, despite receiving no further external input and having a passive decay rate (δ^+) of 0.5. The off-channel activation affecting the new target is, in addition, supported by some residual strength, at the onset of the probe, in the link from the distractor color (green) to the target (as previously noted, in the IP condition, the

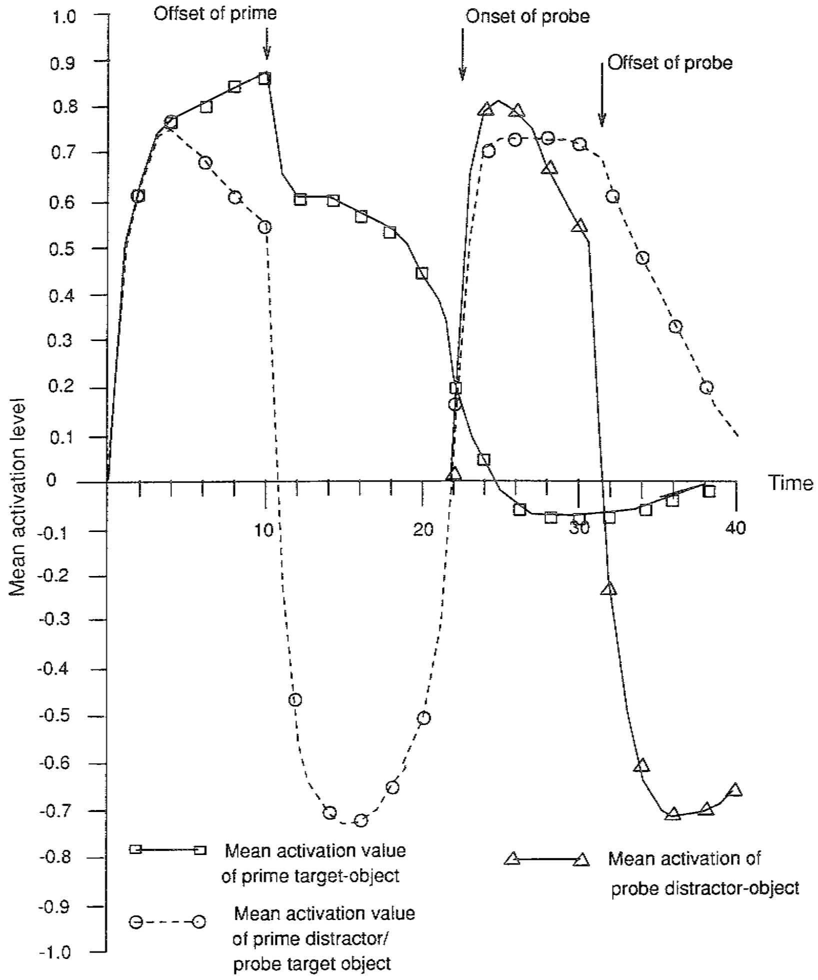

Figure 15

Full simulation of a negative priming trial with both a prime and a probe display. This figure shows the mean activation of object assembly nodes in the IP (ignored prime) condition, that is, where the target in the probe is the previously ignored distractor. In this condition, selection of the target in the probe trial is clearly impaired relative to both the prime trial and to the selection of a novel probe (control condition) shown in the Figure 14.

probe target has changed color from green to red in the probe trial). The idea that negative priming might involve interference in the formation of what we call object assemblies has been previously put forward by Allport, Tipper, and Chmiel (1985).

Thus the model claims that for brief stimulus exposures, negative priming in the IP condition involves both increased interference from the distractor while the display is on (compared to the NP condition) and in addition a "weaker" (more rapidly fading) internal representation of the previously ignored target. Given longer exposure this "weakness" is eventually overcome. Our account of the negative priming effect is thus that the suppressed activation of an ignored distractor found at the offset of a display results in a later response to that distractor, suffering greater interference from copresent objects. This emphasis on the role of interference is very important because it suggests that in the ignored prime condition, negative priming might not be observed if there is no distractor present in the probe stimulus, because there would then be no interference. As is clear from Figure 15, the fact that the representation of the ignored distractor is suppressed at probe onset does not actually prevent that representation from rapidly reaching a significant activation level. What is slowing coherent response binding is the difficulty in selecting the previous distractor as the new target, which requires the suppression of a probe distractor that has an initial activation advantage.

This effect of the distractor can best be shown by comparing the activation of competing categorical response bindings when there either is or is not a distractor present in the probe display in both the IP and NP conditions. Figure 16 shows the difference in the activations of units representing the target and distractor categories during the probe display for each of four conditions: ignored prime, with and without distractor (IP +D, IP −D), and novel probe, with and without distractor (NP +D, NP −D).

Positive values of the curves indicate that the correct (target) category is more highly active. Negative values indicate that the incorrect (distractor) category is more highly active. As can be seen, there is a considerable difference in activation in both IP and NP conditions depending on whether or not a distractor is present. The only time when the incorrect categorization decision shows an advantage is in the IP +D condition. In the absence of a distractor, correct categorization in the IP condition is rapid and effective, though there is still some lag relative to the novel probe condition.

Allport et al. (1985, Experiment 9; see also Lowe, 1979) tested the four probe conditions simulated here and found that in the absence of a probe distractor, RTs in the ignored prime condition were faster than to both the IP +D condition *and* the NP +D condition, in accordance with the model. They also found, however, that the IP −D probe produced faster responding than the corresponding NP −D probe; that is, in the absence of a distractor there was *facilitation* of the ignored prime relative to the novel probe, which we have not replicated in the model (the

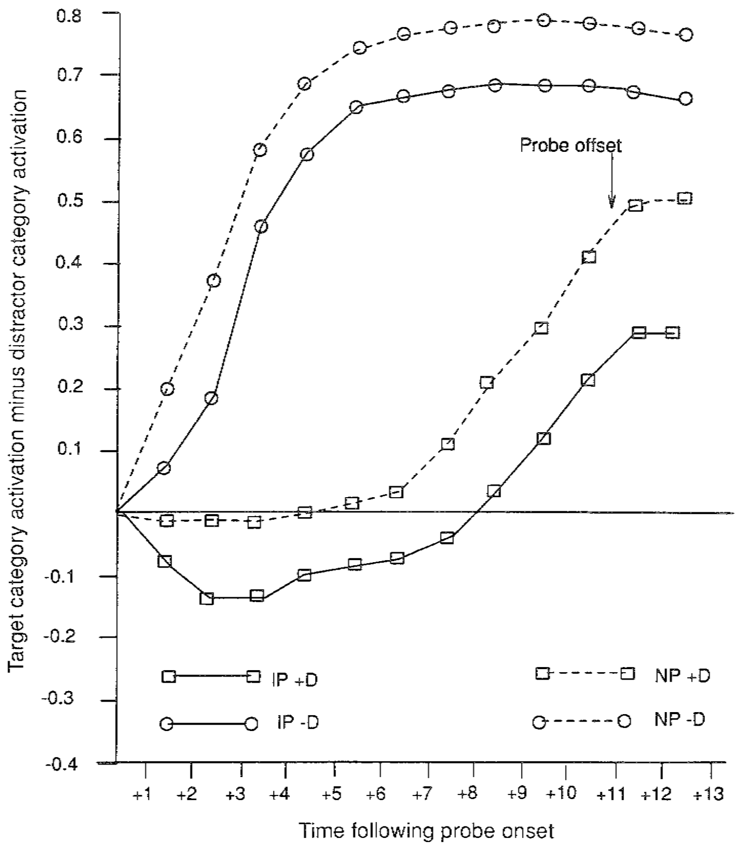

Figure 16

Negative priming as the effect of increased interference in response binding from the distractor during the probe. Each curve shows the difference in activation between the unit in the response binding field representing the correct (target-matching) category and the unit representing the incorrect (distractor) category during presentation of the probe. The four conditions are IP +D: ignored prime with distractor, IP −D: ignored prime without distractor, NP +D: novel probe with distractor, and NP −D: novel probe without distractor. In the ignored prime trials, the to-be-attended item was the distractor on the previous (prime) trial.

NP −D curve in Figure 16 shows the fastest rise and highest asymptote, which we take to be negatively correlated with RT). Thus our model gives the correct ordinal positions for three of the four conditions, but cannot account for the facilitation of the ignored prime (relative to the NP −D condition) in the absence of a distractor. It seems possible (as is argued by Allport et al., 1985) that this facilitation is produced by the prior encoding of the IP −D target in the prime, allied to the fact that in the absence of the distractor, coupling of noncategorical (e.g., color) features with categorical features is not necessary for selection and identification. Responses can therefore be made simply on the basis of the facilitated categorical features. However, as mentioned earlier, we have not tried to account for facilitatory effects generally, as we believe them to be due to different mechanisms.

Thus our model does not include any perceptual facilitation in the formation of an object assembly due to prior encoding of that object. There is nothing in the model, however, which rules out such facilitatory processes, and it seems clear that this explanation for the facilitation of ignored primes could be implemented within our current scheme, for instance by requiring the links in an object assembly to build up over time. Prior formation of links (during the prime) could then facilitate their (re-)formation during the probe. Alternatively, as noted earlier (Figure 9), our model shows activation of response bindings related to the distractor which could form the basis of a small positive priming effect. We should add though that this issue is a little confused empirically. For instance, Tipper et al. (1990) found no priming effects when the probe had no distractor. Others (e.g., Yee, 1991) have found negative priming in the absence of probe distractors. This suggests that whatever the source of facilitatory priming due to ignored distractors, it is fairly weak and probably varies with the task.

We have carried out other simulations with the model investigating, for instance, the effects of parameter variance on the timecourse of the negative priming effect and the relation between efficiency of selection and degree of negative priming. The model also shows semantic spread of negative priming if ignored prime and probe target share nodes (see the preceding discussion of similarity effects in interference). At present, we are testing novel predictions derived from such simulations. Although we anticipate that the model will not remain unchanged in the light of further investigation, we believe that in its current form, it has provided a useful first attempt, suggesting new experiments and providing a much more detailed (and demonstrably effective) theory of selection and negative priming than has hitherto been available.

During work with the computer model, we noted that the rebound effect which the opponent mechanism generates is not confined to ignored inputs, but can happen to selected inputs also, if the internal target field representation suddenly changes. This generates a mismatch with the still activated, previously selected item, causing it to inhibit itself. This is an adaptively useful property, allowing rapid attention switching by automatically deselecting previously facilitated items and preventing perseverative responding. Given the apparent utility of such a process (provided for free by the model), we wondered whether there was any evidence for it in attention switching and we were put in mind of the phenomenon of inhibition of return (IOR; Posner & Cohen, 1984). We consider in the next section whether a mechanism essentially identical to the one we propose can also account for IOR.

C. Inhibition of Return

In this section, we briefly discuss another attentional phenomenon, inhibition of return (Posner & Cohen, 1984). Whereas the negative priming effect appears to be implicated in voluntary selective attention (endogenous selection), IOR arises in the context of the orienting response, the "grabbing" of attention by an external event (exogenous selection). Our main concern here will be to consider whether the IOR phenomenon can be accounted for by mechanisms similar to those we propose in our account of negative priming, and consequently whether endogenous and exogenous selection processes may be related.

In the IOR paradigm, subjects fixating a central point have their attention peripherally cued. If a target subsequently appears in the cued location within about 100 msec, RTs to the target are facilitated with respect to targets at uncued locations. Paradoxically, at longer SOAs of 300–500 msec, RTs to targets at cued locations are retarded by comparison to uncued targets (Maylor, 1985). This phase represents the inhibition of return. Posner and Cohen (1984) argued that the facilitation was due to a short-lived covert orienting response to the cued location, that is, the involuntary capturing of attention by an unexpected external event. The inhibitory effect was also found for overt orienting and appeared to require that after orienting, attention had to be removed from the cued location. According to the authors: "[I]f attention is not drawn away from a cued location, no net inhibition is found" (p. 541). They noted that the inhibitory component apparently occurs automatically (not as the result of a conscious strategy) and argued that it was independent of the facilitatory effect and possibly due to events in sensory pathways rather than to the orienting response per se.

A somewhat different view is put forward by Maylor (1985), who argues that the facilitatory and inhibitory components are not due to different processes, but rather reflect different facets of the orienting reaction. This conclusion is supported by a number of empirical studies carried out by Maylor. For instance, in one experiment (Maylor, 1985, Experiment 2) it was found that the facilitation effect was completely abolished if the subject needed to make an overt orienting response to another stimulus at the moment the cue appeared (thus preventing orienting to the cue). In this case, inhibition disappeared also, apparently being linked to the occurrence of orienting (and thus of facilitation). In addition to this, it was found (Maylor, 1985, Experiment 3) that when cueing of two locations simultaneously occurred, then facilitation was approximately halved compared to cueing to a single location. At the same time, the amount of inhibition found was also decreased in the double-cue condition, again by approximately one half. Maylor concluded that the inhibitory component, as well as the facilitatory, is thus "dependent on externally controlled orienting," and acts in tandem with the facilitation.

A recent development in this paradigm is the discovery by Tipper, Driver, and Weaver (1991) that IOR can occur not just for locations but for whole objects. These authors employed a design in which subjects' attention was cued to one of two peripheral *moving* objects (squares). Following cueing, subjects' attention was recalled to a central point. On two-thirds of the trials, a probe then appeared in one of the two squares, either the previously cued or the uncued (both of which would have moved). Stimulus onset asynchrony (SOA) between the initial peripheral cue and the probe was equiprobably 430 msec or 695 msec, and subjects had to press a key as soon as they noticed the probe. The important comparison was for the probe detection RT in the cued and uncued conditions, with slower RTs in the cued condition indicating inhibition of return to the object (which would have moved since the presentation of the peripheral cue). This result was obtained for both SOAs, indicating that the inhibitory component of the IOR phenomenon can be linked to whole objects rather than simply locations. (Unfortunately, the experiments reported did not look for the facilitatory component at shorter SOAs.) It had been previously argued that IOR is useful in a visual search through static scenes, inhibiting reexamination of already searched locations. Tipper et al. (1991) argue that "tagging fixed spatial coordinates would not permit efficient search through the dynamic scenes of everyday life, in which mobile objects' locations continually change" (p. 290) and thus that object-centered IOR is highly adaptive.

Given the preceding findings, a reasonable theory of IOR would need to account for at least the following:

1. The coupling of the facilitatory and inhibitory components of IOR, such that the inhibitory component manages to "hit" just what was previously facilitated.
2. The timecourse of the phenomenon, such that the facilitation disappears with a shift of attention and the inhibition takes over. What turns off the facilitation and turns on the inhibition?
3. The apparent positive correlations of the strength of the two components found by Maylor (1985), described earlier.

We show here how a plausible extension of our opponent-based selection mechanism to orienting can account for these findings. To do this we need first to present a functional model of what happens during orienting. Our proposed model is shown in Figure 17. The cueing signal is detected by an orienting subsystem (which we propose works in dynamic, competitive interaction with the top-down attentional system in our selection model). The detection of the cue and activation of orienting sets up a representation of an orientation target, analogous to the top-down target in the selection model. (This exogenously generated target will typically suppress endogenous targets, diverting attention from ongoing goal-directed activity.) We propose that this target interacts with the cue representation in the same way as in our selection model. The cue representation (which may, of course, be an object) is composed of linked opponent circuits, and the interaction of cue and target in the match/mismatch system generates a match signal causing facilitation in the manner described earlier. (This matching would permit the organism to know when orienting is complete.) We propose that the subsequent switching away of attention involves the quenching of the activation of the previous (externally derived) orientation target and its replacement with another target (specified by an internal plan or a new external stimulus, for instance). The activation in the cue representation will persist, however (due to the previous facilitation, or to the

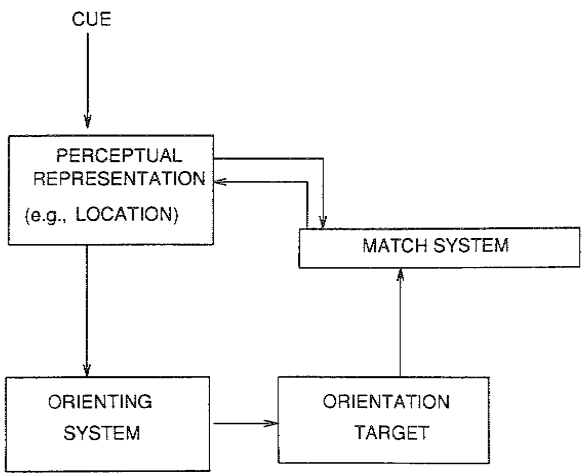

Figure 17
A functional model of the orienting response. We propose that cue detection causes the orienting system to set up an internal representation of the cue location, which then acts as a target. The presence of the internal target leads to facilitation of the cue location by feedback via the match system. The internal architecture and interactions of the target, match, and perceptual systems in the figure are identical to the target, match, and object fields (respectively) of our selection model.

fact that the corresponding object is still visible), but will now generate a mismatch signal from the match field (due to the change of target). This instigates an inhibitory rebound in the representation of the cued input, as occurs to the ignored distractor in our selective attention model. Such a mechanism would clearly be of adaptive value, allowing the organism to rapidly and specifically suppress previously attended (and potentiated) stimuli, thereby preventing interference and perseverative responding. (Regarding perseveration, we note in passing that the need to actively suppress highly activated representations in the control of serially ordered behavior is postulated in the "competitive queueing" models of Houghton, 1990, and Burgess & Hitch, 1992. Both these models use an opponent-type mechanism to achieve this, whereby a sequence element generates inhibitory feedback onto itself. See Houghton, 1993, for discussion of the general need for such control mechanisms in neural networks.)

We now illustrate this dynamic in a number of simulations in which we adapt our selection model to the orienting response, with minimal changes. In terms of Figure 17, the target, match, and perceptual representation systems are identical in internal structure and interactions to the target, match, and object fields of our selective attention model. We simply add the assumption that the orienting system, on being activated by a cue, creates an orientation target, represented as activity in a target node. All the parameters used in the following simulations are kept constant and are identical to those used in the previous simulations.

A simulation of the basic dynamical pattern we propose for IOR is shown in Figure 18. The curve shown represents the activation level of a location node in the object field which is activated by an external cue. On orienting to this location (realized by the instantiation of the cue location as a target), its activation is boosted. At t = 5, attention is switched away (coincident, in this case, with cue offset). This attention switching is implemented in the simulation by the replacement of the target location with another target (representing, say, the location of a central fixation point). The cue representation now no longer matches the target and quickly suffers an inhibitory rebound, in essentially the same manner as do the ignored distractors in our selection model.

As previously noted, Posner and Cohen (1984) found that the inhibitory component of IOR only occurs after attention is switched from the cued location. Figure 19 illustrates that this is also true for our model and that the effect is not contingent on cue offset coinciding with an attentional shift. In this case, attention is maintained to the cued location after cue offset (leading to facilitation). Only on attentional shift (t = 14) is the inhibitory component initiated.

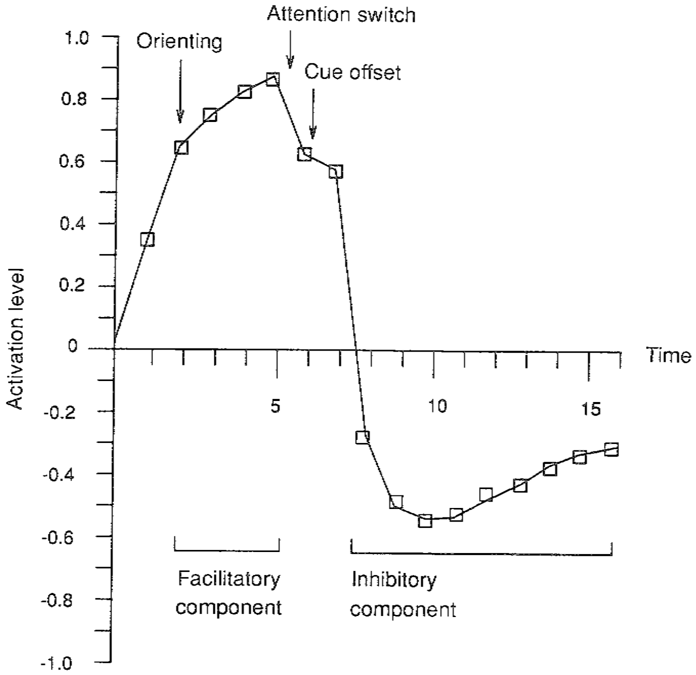

Figure 18

Simulation of the two components of the inhibition of return effect. The curve represents the activation of a single node representing the location of an external cue. Orienting (realized as the establishment of the internal target) takes place at t = 2, leading to facilitation. Attention switching (followed rapidly by cue offset in this case) at t = 5 removes the facilitation and initiates an inhibitory rebound in the location opponent circuit. This suppresses activation of the previously facilitated location.

It was proposed that a model of IOR should account for Maylor's finding that (in a multiple cueing task) degree of facilitation was positively correlated with degree of inhibition (less facilitation, less inhibition). In attempting to simulate this phenomenon, the question arises as to how one should represent the effect of having multiple cues in the model. We propose here that having multiple cues leads to each cue becoming more weakly represented as an orientation target. This weakened activation might, for instance, be due to competitive interactions within the orienting system. In the previous simulations, the single orientation target was given an activation value of 1 (as in all the other simulations reported here). The strength of activation of the internal target affects the degree of facilitation of matching inputs by virtue of the fact that the strength of the match signal is computed as the product of the cue signal and the target signal [Eq. (9)]. This signal, in turn, affects the degree to which the cue opponent circuit is pushed over into excitatory mode. The greater the target activation level then, the greater

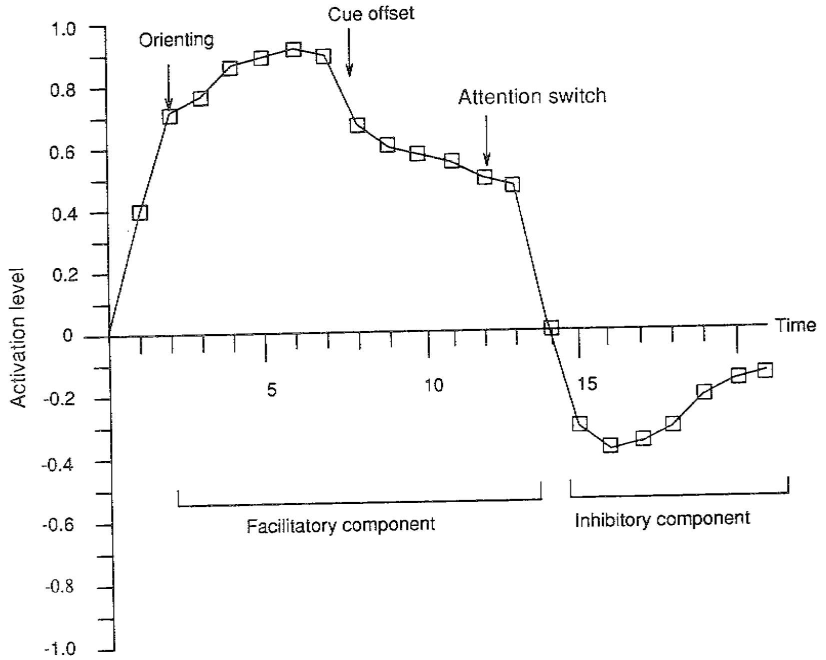

Figure 19
This figure illustrates that in the model the inhibitory rebound is initiated by attention switching, not cue offset. This is in agreement with the findings of Posner and Cohen (1984).

should be the facilitation. The question then arises whether under these circumstances increased (or decreased) cue facilitation in the model will lead to increased (or decreased) inhibition of return. Figure 20 shows that we get a positive correlation.

In this simulation, the model was run with target activation levels increasing from 0.2 to 1 in steps of 0.1. The timing of cue onset, offset, and attention switching were the same as for the simulation in Figure 18. For each run of the model, the maximum and minimum values of the cue node activation level were recorded. The maximum value was taken to indicate degree of facilitation, and the minimum value, degree of inhibition. Each vertical line on the graph joins the maximum and minimum activation values achieved for a given target activation value (shown below the line). As can be seen, both facilitation and inhibition increase in tandem as target activation strength increases. We thus attempt to account for the data by claiming that multiple cueing leads to weaker target activation (weaker cue activation per se is also possible). This theory could be independently tested if we can equate the level of target activation with other psychological variables, such as degree of focusing or concentration. The model predicts a similar pattern of results for single cues if, say, attention is divided (or concentration is low).

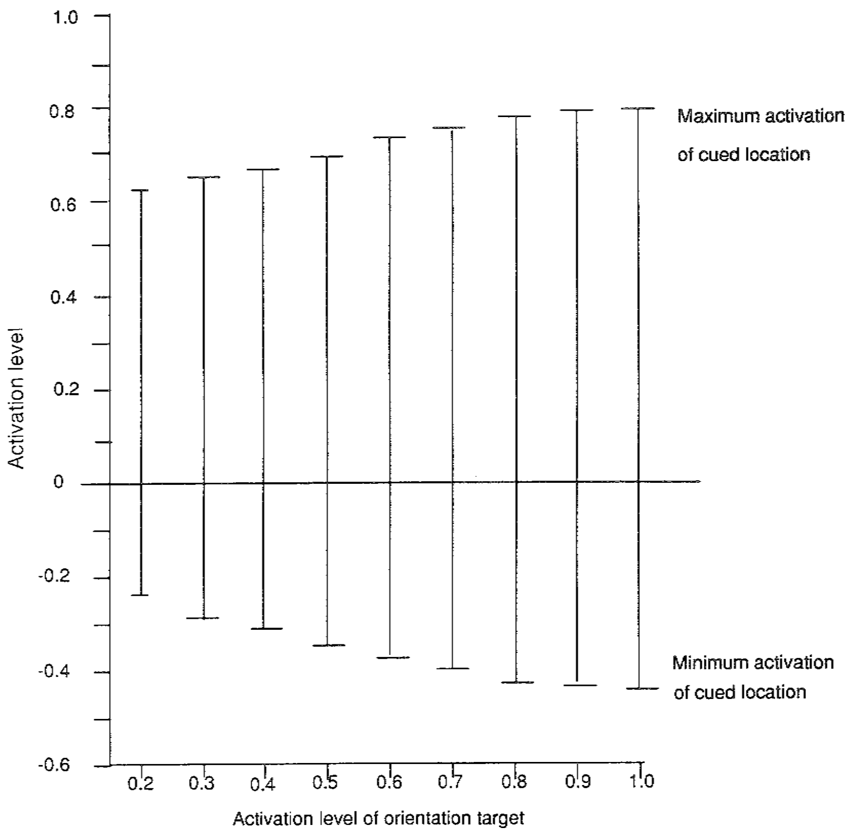

Figure 20
In simulations of inhibition of return (IOR), greater activation of the orientation target node (derived exogenously from cueing), shown along the *x*-axis, leads to greater facilitation of the cued location, followed by greater IOR to that location, as indexed by the degree of suppression of the sensory representation of the location. Each vertical line in the graph joins points representing maximum and minimum activation values attained by a cued location mode for a given level of internal target node activation.

The simulations described show that the model provides answers to the theoretical questions about IOR we asked earlier. It explains the coupling of the facilitatory and inhibitory components (how inhibition "knows" what has previously been facilitated) by having the sources of both components joined in the opponent mechanism—it is the facilitation itself which primarily drives the inhibitory feedback. The timecourse and the dependence on attention switching for the release of inhibition is explained by the fact that, in the model, the switching of the opponent feedback circuit from net excitation to net inhibition is dependent on the generation of match/mismatch signals. The switch of attention means that the cue representation no longer matches the target and it effectively quenches itself. This account is also in line with Posner and Cohen's claim that the inhibitory component is automatic. The positive correlation found between degree of facilitation and inhibition (in multiple cueing) is explained on the assumption that it is reasonable to equate multiple cues with weaker internal activation of each one (other things being equal). Weaker targets lead to less facilitation, and thereby to a weaker inhibitory rebound.

In conclusion, our model provides a concrete demonstration that the IOR phenomenon (and hence orienting) might involve mechanisms very similar to those involved in voluntary selective attention and negative priming. In voluntary selection, the mismatch-driven opponent mechanism actively deselects nontarget items while maintaining their activation at an equilibrium value above resting levels. Offset of external inputs to a deselected item throws the system out of equilibrium, causing it to experience an inhibitory rebound manifest as negative priming. In the case of inhibition of return, the rebound is caused by an internal switch of target specification, which leads a previously selected item to automatically deselect itself, preventing it from interfering with subsequent processing. We believe this account has sufficient theoretical elegance and economy to deserve active investigation.

V. CONCLUSIONS

This chapter advances a theory of selective attention formulated at a number of levels. At the functional level, we propose that selective attention facilitates the organism's maintenance of its goal-directed behavior by gating the flow of perceptual information into response systems (conceived to encompass both action and thought), emphasizing goal-relevant information, and backgrounding irrelevant or contextual information. In neuropsychological terms, we suggest this involves the action of the prefrontal lobes in gating the flow of activation from posterior (perceptual) systems through to frontal motor planning and execution systems. In terms of mechanisms, we propose that attentional gating requires the maintenance of internal target specifications that are matched against high-level perceptual representations. Matched representations are facilitated, and nonmatching inhibited. The precise mechanism we propose for this operation allows efficient selection over the whole dynamic range of the processing substrate and implements automatic gain control of inhibitory signals, so that the strength of the inhibition continually adapts to the strength of the to-be-ignored inputs. This self-regulating feedback mechanism allows nontarget stimuli to find automatically an equilibrium activation level below that of targets, but above resting levels. The opponent mechanism realizing these properties exhibits rebound behavior, which we use to account for the phenomenon of negative priming, bringing our model into contact with established experimental paradigms. In addition, the same rebound behavior, triggered somewhat differently, allows for rapid attention switching and provides an account of inhibition of return. In conclusion, the model contains mechanisms motivated primarily on functional grounds (e.g., the provision of self-regulating inhibitory feedback), which turn out to have certain nonobvious properties enabling us to account for a range of data not previously explained in any theoretically precise or coherent manner. Future work will concentrate on further testing of the model's predictions, as well as investigate variants of and alternatives to the model itself.

REFERENCES

Allport, D. A., Tipper, S. P., & Chmiel, N. (1985). Perceptual integration and postcategorical filtering. In M. I. Posner and O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 107–132). Hillsdale, NJ: Lawrence Erlbaum.

Arbib, M. A., (1990). Programs, schemas, and neural networks for control of hand movements: Beyond the RS framework. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 111–138). Hillsdale, NJ: Lawrence Erlbaum.

Arbib, M. A., Iberall, T., & Lyons, D. (1987). Schemas that integrate vision and touch for hand control. In M. A. Arbib & R. Hanson (Eds.), *Vision, brain and co-operative computation* (pp. 489–510). Cambridge, MA: MIT Press.

Arsten, A. M. F., Segal, D. S., Neville, H. J., Hillyard, S. A., Janowsky, D. S., Judd, L. L., & Bloom, F. E. (1983). Naloxone augments electrophysiological signs of selective attention in man. *Nature, 304*, 725–727.

Beech, A. R., Agar, K., & Baylis, G. C. (1989). Reversing priming while maintaining interference. *Bulletin of the Psychonomic Society, 27*, 553–555.

Beech, A. R., Baylis, G. C., Smithson, P., & Claridge, G. (1989). Individual differences in schizotypy as reflected in measures of cognitive inhibition. *British Journal of Clinical Psychology, 28*, 117–129.

Beech, A. R., Baylis, G. C., Tipper, S. P., McManus, D., & Agar, K. (1991). Individual differences in cognitive processes: Towards an explanation of schizophrenic symptomtology. *British Journal of Psychology, 82*, 479–489.

Beech, A. R., & Claridge, G. (1987). Individual differences in negative priming: Relations with schizotypal personality traits. *British Journal of Psychology, 78*, 349–356.

Beech, A. R., Powell, T. J., McWilliams, J., & Claridge, G. S. (1989). Evidence of reduced "cognitive inhibition" in schizophrenia. *British Journal of Clinical Psychology, 28*, 110–116.

Beech, A. R., Powell, T. J., McWilliams, J., & Claridge, G. S. (1990). The effect of a small dose of chlorpromazine on a measure of cognitive inhibition. *Personality and Individual Differences, 11*, 1141–1145.

Benoit, G., Fortran, L., Lemelin, S., LaPlante, L., Thomas, J., & Everett, J. (1992). L'attention selective dans la depression majeure: Ralentissement clinique et inhibition cognitive. *Canadian Journal of Psychology, 46*, 41–52.

Biederman, I. (1972). Perceiving real world scenes. *Science, 177*, 77–80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1981). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143–177.

Bjork, E. L., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review, 84*, 472–484.

Briand, K. (in preparation). *Selective attention to global and local structure of objects: Converging measures of nontarget processing.*

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

Broadbent, D. E. (1970). Stimulus and response set: Two kinds of selective attention. In D. I. Mostotsky (Ed.), *Attention: Contemporary theories and analysis.* New York: Appleton-Century-Crofts.

Burgess, N., & Hitch, G. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31*, 429–460.

Curio, E. (1976). *The ethology of predation.* New York: Springer-Verlag.

Dalrymple-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop color-word test. *Perceptual and Motor Skills, 23*, 1211–1214.

DeSchepper, B. G., Khurana, B., O'Connell, K. M., & Wilson, M. (in preparation). *What is inhibited in negative priming?*

DeSchepper, B., & Treisman, A. (1991). Novel visual shapes in negative priming. *Thirty-second Annual Meeting of the Psychonomic Society, San Francisco.*

Driver, J., & Baylis, G. C. (1993). Cross-modal negative priming and Stroop-like interference in selective attention. *Bulletin of the Psychonomic Society, 31*, 45–48.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87*, 272–300.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Enright, S. J., & Beech, A. R. (1990). Obsessional states: Anxiety disorders or schizotypes? An information processing and personality assessment. *Psychological Medicine, 20*, 621–627.

Eriksen, C. W., Coles, M. G. H., Morris, L. R., & O'Hara, W. P. (1985). An electromyographic examination of response competition. *Bulletin of the Psychonomic Society, 23*, 165–168.

Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception and Psychphysics, 12*, 278–286.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108*, 316–355.

Fuentes, L. J., & Tudela, P. (1992). Semantic processing of foveally and parafoveally presented words in a lexical decision task. *Quarterly Journal of Experimental Psychology, 45A*(2), 299–322.

Fuster, J. M. (1980). *The prefrontal cortex* [2nd ed., 1989]. New York: Raven Press.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99–156.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 245–262.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In F. Plum (Ed.), Handbook of physiology: *Sec. 1.* The nervous system: Vol. V: Higher cortical function (pp. 373–417). Bethesda, MD: American Physiological Society.

Goldman-Rakic, P. S. (1988). Topography of cognition: Parallel distributed networks in primate association cortex. *Annual Review of Neuroscience, 11*, 137–156.

Gray, J. A., Feldon, J., Rawlins, J. N. P., Hemsley, D. R., & Smith, A. D. (1991). The neuropsychology of schizophrenia. *Behavioral and Brain Sciences, 14*, 1–84.

Greenwald, A. C. (1972). Evidence both for perceptual filtering and response suppression for rejected messages in selective attention. *Journal of Experimental Psychology, 94*, 58–67.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review, 87*, 1–51.

Grossberg, S. (1983). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Psychological Review, 89*, 529–572.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 163–169.

Hoffman, J. E., & MacMillan, F. W. (1985). Is semantic priming automatic? In M. I. Posner and O. S. M. Marin (Eds.), *Attention and performance XI.* Hillsdale, NJ: Lawrence Erlbaum.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 287–319). London: Academic Press.

Houghton, G. (1993). Inhibitory control of neurodynamics: Opponent mechanisms in sequencing and selective attention. In M. Oaksford & G. D. A. Brown (Eds.), *Neurodynamics and Psychology.* London: Academic Press.

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

Hung, D. L., & Tzeng, O. J. L. (1989, November). Location-specific inhibitory processes in visual selective attention. *Thirtieth Annual Meeting of the Psychonomic Society,* New Orleans.

Hurvich, L. M., & Jameson, D. (1957). An opponent process theory of color vision. *Psychological Review, 64*, 384–404.

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

Kahneman, D., & Treisman, A. M. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. Davis (Eds.), *Varieties of attention* (pp. 29–61). Orlando, FL: Academic Press.

Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology, 57*, 576–588.

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

Kohonen, T. (1984). *Self-organization and associative memory.* New York: Springer-Verlag.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96*, 101–124.

La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory & Cognition, 16*, 400–410.

La Heij, W., Dirkx, J., & Kramer, P. (1990). Categorical interference and associative priming in picture naming. *British Journal of Psychology, 81*, 511–525.

La Heij, W., Van Der Heijden, A. H. C., & Schreuder, R. (1985). Semantic priming and Stroop-like interference in word naming tasks. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 62–80.

Lowe, D. G. (1979). Strategies, content, and the mechanisms of response inhibition. *Memory & Cognition, 7*, 382–389.

Lowe, D. G. (1985). Further investigations of inhibitory mechanisms in attention. *Memory & Cognition, 13*, 74–80.

Luria, A. R. (1973). *The working brain.* London: Penguin.

Marr, D. (1982). *Vision.* New York: W. H. Freeman.

Maylor, E. A. (1985). Facilitatory and inhibitory components of orienting in visual space. In M. I. Posner & O.S.M. Marin (Eds.), *Attention and performance XI* (pp. 189–204). Hillsdale, NJ: Lawrence Erlbaum.

McDowd, J. M., & Oseas-Kreger, D. M. (1991). Aging, inhibitory processes, and negative priming. *Journal of Gerontology, 46,* 340–345.

McLaren, J. (1989). *The development of selective and sustained attention in normal and attentionally disordered children.* Unpublished doctoral dissertation, Dalhousie University, Halifax, Nova Scotia.

McLaren, J., & Bryson, S. (1988). Sustained and selective attention. *Canadian Psychology, 29,* Abstract No. 530.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London, B298,* 65–95.

Mishkin, M., & Appenzeller, T. (1987). The anatomy of memory. *Scientific American, 256*(6), 62–71.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6,* 414–417.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Mueller, P. M., & Baylis, G. C. (submitted). *Effects of normal aging and Alzheimer's disease on attention: 2. Intertrial priming effects.*

Näätänen, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. *Psychological Bulletin, 92,* 605–640.

Näätänen, R. (1985). Selective attention and stimulus processing: Reflections in event-related potentials, magnetoencephalogram, and regional blood flow. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance* XI (pp. 355–373). Hillsdale, NJ: Lawrence Erlbaum.

Neill, S. R., & Cullen, J. M. (1974). Experiments on whether schooling by their prey affects the hunting behavior of celaphod and fish predators. *Journal of Zoology,* London, *172,* 549–569.

Neill, W. T. (1977). Inhibition and facilitation processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 444–450.

Neill, W. T., & Valdes, L. A. (1992). The persistence of negative priming: Steady-state or decay? *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 565–576.

Neill, W. T., & Westberry, R. L. (1987). Selective attention and the suppression of cognitive noise. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13,* 327–334.

Neumann, E., & DeSchepper, B. G. (1991). Costs and benefits of target activation and distractor inhibition in selective attention. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17,* 1136–1145.

Neumann, E., & DeSchepper, B. G. (1992). An inhibition-based fan effect: Evidence for an active suppression mechanism in selective attention. *Canadian Journal of Psychology, 46,* 11–50.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer and A. F. Sanders (Eds.), *Perspectives on perception and action.* Hillsdale, NJ: Lawrence Erlbaum.

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.), *Conscious and self-regulation: Advances in research:* Vol. IV. New York: Plenum Press.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Picton, T. W., Stuss, D. T., & Marshall, K. C. (1986). Attention and the brain. In S. L. Friedman, K. A. Klivington, & R. W. Peterson (Eds.), *The Brain, Cognition, and Education* (pp. 19–79). New York: Academic Press.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. I., & Cohen, Y. A. (1984). Components of visual orienting. In H. Bouma and D. G. Bouwhuis (Eds.), *Attention and performance* (pp. 531–554). Hillsdale, NJ: Lawrence Erlbaum.

Posner, M. I., & Driver, J. (1992). The neurobiology of selective attention. *Current Opinion in Neurobiology, 2,* 162–169.

Roland, P. E. (1982). Cortical regulation of selective attention in man. A regional cerebral blood flow study. *Journal of Neurophysiology, 48,* 1059–1078.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Vol. 1, Foundations.* Cambridge, MA: MIT Press.

Schull, J. (1979). A conditioned opponent theory of Pavlovian conditioning and habituation. In G. H. Bower (Ed.), *The psychology of learning and motivation:* Vol. 13. New York: Academic Press.

Solomon, R. L., & Corbit, J. D. (1974). An opponent process theory of motivation: 1. Temporal dynamics of affect. *Psychological Review, 81,* 119–145.

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

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes.* New York: Raven Press.

Swerdlow, N. R., & Koob, G. F. (1987). Dopamine, schizophrenia, mania and depression: Toward a unified hypothesis of cortico-striato-pallido-thalamic function. *Behavioral and Brain Sciences, 10,* 197–245.

Taffe, M., Moore, B. O., Tipper, S. P., & Baylis, G. C. (1991). Action-based spatial attention in normal and hippocampal lesioned monkeys. *Society for Neuroscience Abstracts, 18,* 479.9.

Tipper, S. P. (1985). The negative priming effect: Inhibitory effects of ignored primes. *Quarterly Journal of Experimental Psychology, 37A,* 571–590.

Tipper, S. P. (1991). Less attentional selectivity as a result of declining inhibition in older adults. *Bulletin of the Psychonomic Society, 29,* 45–47.

Tipper, S. P., & Baylis, G. C. (1987). Individual differences in selective attention: The relation of priming and interference to cognitive failure. *Personality and Individual Differences, 8,* 667–675.

Tipper, S. P., Bourque, T., Anderson, S., & Brehaut, J. C. (1989). Mechanisms of attention: A developmental study. *Journal of Experimental Child Psychology, 48,* 353–378.

Tipper, S. P., Brehaut, J. C., & Driver, J. (1990). Selection of moving and static objects for the control of spatially directed action. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 492–504.

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *Quarterly Journal of Experimental Psychology, 37A,* 591–611.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words: Evidence for semantic analysis of ignored stimuli. *Memory & Cognition, 16,* 64–70.

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

Tipper, S. P., Lortie, C., & Baylis, G. C. (1992). Selective reaching: Evidence for action-centered attention. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 891–905.

Tipper, S. P., MacQueen, G. M., & Brehaut, J. C. (1988). Negative priming between response modalities: Evidence for the central locus of inhibition in selective attention. *Perception and Psychophysics, 43,* 45–52.

Tipper, S. P., & McLaren, J. (1990). Evidence for efficient visual selective attention in children. In J. T. Enns (Ed.), *The development of attention: Research and theory.* Amsterdam: Elsevier.

Tipper, S. P., Weaver, B., Cameron, S., Brehaut, J. C., & Bastedo, J. (1991). Inhibitory mechanisms of attention in identification and localization tasks: Timecourse and disruption. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 681–692.

Tipper, S. P., Weaver, B., & Houghton, G. (in press). *Behavioral goals determine excitation–inhibition states of the internal representations of unattended objects. Quarterly Journal of Experimental Psychology.*

Tipper, S. P., Weaver, B., Kirkpatrick, J., & Lewis, S. (1991). Inhibitory mechanisms of attention: Locus, stability, and relationship with distractor interference effects. *British Journal of Psychology, 82,* 507–520.

Treisman, A. M. (1986). Features and objects in visual processing. *Scientific American, 255,* 114–125.

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

Underwood, G. (1980). Attention and the nonselective lexical access of ambiguous words. *Canadian Journal of Psychology, 34,* 72–76.

Van Der Heijden, A. H. C. (1981). *Short term visual forgetting.* London: Routledge and Keegan Paul.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms.* Hillsdale, NJ: Erlbaum.

Wundt, W. (1904). *Principles of physiological psychology* (5th ed.) [originally published 1874]. Fifth edition, E. Titchener (Ed.). New York: Macmillan.

Yee, P. L. (1991). Semantic inhibition of ignored words during a figure classification task. *Quarterly Journal of Experimental Psychology, 43A,* 127–153.

3

Categories of Cognitive Inhibition with Reference to Attention

Raymond M. Klein and Tracy L. Taylor

I. INTRODUCTION

In contemporary psychology there is great interest in the interdisciplinary wedding of the cognitive and neural sciences.[1] In this burgeoning field of cognitive neuroscience, information from cognitive science informs the study of the ways in which cognitive computations may be implemented in the neural machinery, and, in a similar manner, knowledge of neural functioning helps to develop and constrain models of cognitive functioning. The potential for symbiosis between cognitive and neural science is particularly evident in the study of inhibition where, on the one hand, neural mechanisms provide a terminology and insight for understanding possible mechanisms of cognitive inhibition and, on the other hand, behavior-based models of cognitive functioning imply the need for inhibitory circuitry. Distinct from neural inhibition, cognitive inhibition refers to a hypothetical construct whose effects within the information processing stream are inferred to exist on the basis of observable human behavior. By reducing or blocking the

---

[1] The excitement of this occasion has generated interest in models of cognitive processes in which vast arrays of pseudoneural elements are richly interconnected in a neo-Hebbian network of excitatory and inhibitory links in order to simulate some aspect of human behavior. Whether one regards this approach as one method for developing neuropsychological theories or as a pretheoretical assumption/belief about how such theories ought to be expressed, neoconnectionism is and will continue to be a fruitful approach (cf. Houghton & Tipper, this volume).