

Perceptual inhibition of expected inputs: The key that opens closed minds

WILLIAM A. JOHNSTON and KEVIN J. HAWLEY
University of Utah, Salt Lake City, Utah

The mind appears to be biased simultaneously toward both expected and unexpected inputs. For example, familiar scenes are usually perceived more readily than novel scenes, indicating the former bias, but a single novel object sometimes pops out from a familiar field, indicating the latter bias. A diverse literature and a computational model converge on the following resolution to this paradox: The former bias is conceptually driven and actually suppresses data-driven processing of expected inputs; in turn, this suppression disinhibits data-driven processing of unexpected inputs, yielding the latter bias. Evidence for suppressed data-driven processing of expected inputs is drawn from studies of perceptual habituation, semantic satiation, memory inhibition, inhibition of return, repetition blindness, primed inhibition, the word-inferiority effect, registration without learning, and both expert- and schema-based inhibitory effects. Evidence for enhanced data-driven processing of unexpected inputs is drawn from studies of the orienting response, mismatch negativity, memory facilitation, both expert- and schema-based facilitatory effects, and perceptual popout. The model, called *mismatch theory*, incorporates inhibitory and facilitatory perceptual dynamics and is found to simulate the opposing biases. Implications of mismatch theory for perceptual phenomenology, dynamic systems theory, mental health, and individual differences are also discussed.

In a review of some of her research on perception and memory, Treisman (1992) concluded that

by creating accumulated traces of past perceptual objects or events, the world molds our minds to recreate earlier experiences. At the same time, we retain an impressive capacity also to represent any new object that fails to find its match in our prior assembly of stored tokens. (p. 874)

This observation echoes what Grossberg (1987) calls the *stability-plasticity dilemma*. How can the mind be molded to familiar environments and still remain vigilant for change? We suggest that the opposing biases toward both what is most expected and what is least expected are among the most adaptive and revealing features of the mind. Part 1 of this paper presents a paradigmatic example of opposing mental biases in the form of our own prior work on novel popout. Part 2 outlines in general terms a possible explanation of novel popout and resolution of the stability-plasticity dilemma, called *mismatch theory*. Part 3 draws together diverse lines of empirical support for the key assumptions of mismatch theory. Part 4 de-

scribes an explicit computational instantiation of mismatch theory and its power to simulate novel popout. Finally, Part 5 discusses various implications of the dynamic interplay between opposing mental biases.

PART 1 A Case Study of Opposing Mental Biases

In several prior studies, we have investigated what captures attention on a first glance at a scene (see, e.g., Hawley, Johnston, & Farnham, 1994; Johnston, Hawley, & Farnham, 1993; Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990). In a typical study, observers are given 200-msec glimpses of four-object arrays, each of which is followed immediately by one of the objects as a probe. The observer indicates, or guesses, the location that the probe had occupied in the preceding array. Some of the objects, called *familiar*, appear many times across a long series of trials; others, called *novel*, appear only once. Our standard design includes three types of array: *all-novel*, in which only novel objects appear, *all-familiar*, in which only familiar objects appear, and *one-novel*, in which a single novel object appears in a field of three familiar objects. The typical pattern of results includes three basic effects, called *baseline*, *novel popout*, and *familiar sink-in*. These effects are exemplified in the left panel of Figure 1, which summarizes the findings of Experiment 4 of Johnston et al. (1990). The baseline effect is defined by the higher accuracy of localization for all-familiar arrays than for all-novel arrays. This effect illustrates the mental bias toward expected inputs. Novel popout is defined by the higher accuracy of localization for novel ob-

The preparation of this paper was supported by Grant F49620-92-j-0473 from the Air Force Office of Scientific Research to William A. Johnston. Some of the ideas for this paper were stimulated by discussions that William Johnston had with José María Ruiz-Vargas at the University of Madrid. We are grateful to Veronica Dark, Jim Farnham, Douglas Hintzman, David Sanbonmatsu, and David Strayer for their contributions and suggestions at various stages in the evolution of this paper. We also appreciate the guidance of Jim Neely, the action editor for the paper, and of Nelson Cowan, Tram Neill, and Endel Tulving, the consulting editors. Requests for reprints should be sent to W. A. Johnston, Department of Psychology, University of Utah, Salt Lake City, UT 84112.

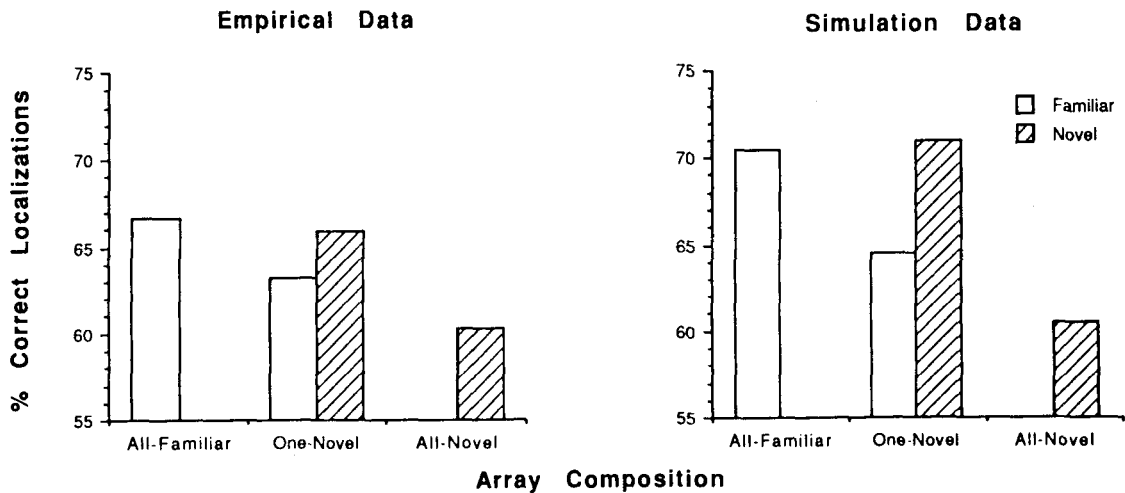


Figure 1. Accuracy of localization for novel and familiar words in all-familiar, one-novel, and all-novel arrays. Empirical data are summarized in the left panel, and simulation data are summarized in the right panel.

jects in one-novel arrays than for those in all-novel arrays, and familiar sink-in is defined by the lower accuracy for the familiar objects in one-novel arrays than for those in all-familiar arrays. These effects illustrate the mental bias toward unexpected inputs—in particular, those that arise from otherwise familiar fields. As Figure 1 illustrates, these effects can at times be of sufficient magnitude so that the novel intrusions are more localizable than are familiar field objects. However, this effect is dependent on the magnitude of the baseline effect and is not always observed (see, e.g., Hawley et al., 1994).

In mapping the boundaries of the three basic effects, we have found them to be relatively independent of type of measure (accuracy of identification vs. localization), duration of array exposure (33–400 msec), type of object (words vs. nonsense objects), spatial predictability of the familiar field objects (variable vs. fixed locations), number of prior occurrences (from 15 to 200) of the familiar objects, and speed/accuracy tradeoffs. However, the effects require that the familiar field objects be collectively, rather than just individually, familiar and contextually or episodically, rather than just preexperimentally, interassociated. For example, the effects were not observed when the field objects either had appeared many times in the context of the experiment but never together or were semantically related but had not appeared previously in the context of the experiment (but see DeWitt, 1994, discussed below). Finally, novel popout is not attributable to a familiarity-based figure-ground contrast effect because a complementary familiar popout effect was not observed when a lone familiar object was embedded in a field of novel objects.

PART 2

Mismatch Theory of Novel Popout

How can the baseline advantage of familiar over novel scenes be reconciled with novel popout and familiar sink-

in? How can the mind both tune into familiar environments and remain open to change? Mismatch theory attributes these opposing functions to different levels of perceptual processing—namely, conceptually driven and data-driven processing, respectively.¹

The basic idea is that because the mind already knows its familiar habitats, it does not waste time and energy reaffirming their details. Instead, on encountering such an environment, the mind relies on its knowledge of that environment (i.e., conceptually driven processing) to perceive it and efficiently navigate the body through it. After an initial bottom-up triggering of the appropriate conceptual processing, data-driven processing is turned down. A by-product of this dampening of the data-driven processing of expected inputs is an accentuation of the data-driven processing of any unexpected inputs in their midst. The computational details of these processes are explicated in Part 4. Suffice it to note here that the inhibited data-driven processing of expected inputs effects a dynamic balance between the opposing mental biases without the aid of special novelty detectors, resource managers, executives, or other homunculi.

In its global form, mismatch theory is similar to the dishabituation theory of Sokolov (1963) and the adaptive resonance theory of Grossberg (1987). All three theories assume that bottom-up processing is suppressed for inputs that match expectation and accentuated for those that do not. However, mismatch theory is fundamentally incompatible with what may be the most widely accepted class of theory of perceptual processing, the prototypical exemplar of which is the interactive-activation model of McClelland and Rumelhart (1981). As with mismatch theory, interactive-activation models assume that perceptual processing of familiar environments is both data-driven by physical inputs and conceptually driven by active schemata. Moreover, as with mismatch theory, these models assume that the two levels or forms of processing are interactive and interdependent. However, in direct contrast to

mismatch theory, these theories assume that the two levels are mutually supportive such that data-driven processing is enhanced for inputs that match expectations and inhibited for inputs that do not. Indeed, in comparison with these models, mismatch theory could be referred to as an interactive-*inhibition* model.

In brief, mismatch theory attributes the opposing mental biases toward both expected and unexpected inputs to different levels of processing. The bias toward expected inputs and familiar environments is attributed to enhanced conceptually driven processing (e.g., spreading excitation across an upper layer of conceptual nodes), and the bias toward unexpected inputs is attributed to enhanced data-driven processing (e.g., bottom-up excitation from a lower tier of feature nodes). The key to these simultaneous biases and the dynamic balance between them is the top-down inhibition of expected inputs.

PART 3 Supporting Evidence

In this section, we review a wide assortment of empirical evidence for the hypothesis that data-driven processing is both inhibited for expected inputs and accentuated for unexpected inputs. This review is admittedly selective and is intended more to lend credence to the hypothesis than to exhaust the potentially relevant literature. The extent to which the various phenomena reviewed are based on common processes, especially those constituting mismatch theory, remains to be investigated.

Perceptual Inhibition of Expected Inputs

The baseline effect in our studies of novel popout is buttressed by a host of evidence for the normally fluent perceptual processing of familiar objects and scenes (see, e.g., Biederman, Glass, & Stacy, 1973; Jacoby & Dallas, 1981; Shiffrin & Schneider, 1977). The research reviewed below provides converging evidence for the idea that these facilitative effects of expectancy conceal an underlying inhibitory effect. This research includes studies of perceptual habituation, semantic satiation, memory inhibition, inhibition of return, repetition blindness, primed inhibition, the word-inferiority effect, registration without learning, expertise-based inhibition, and schema-based inhibition.

Perceptual habituation. The habituation literature attests that neural and perceptual responsivity decline with repeated exposures to the same object (see, e.g., Sokolov, 1963). Yet, studies of perceptual memory show that repetition enhances the perceptibility of objects (see, e.g., Jacoby & Dallas, 1981). A possible reconciliation of the phenomena of perceptual habituation and enhanced perceptual memory was developed and tested by Kraut and Smothergill (1978). They suggested that stimulus repetition has two opposing effects: an alertness decrement, which is relatively short-lasting, and a conceptual-encoding increment, which is relatively long-lasting. The former effect exemplifies the inhibitory effect of expectation hy-

pothesized above; the latter effect exemplifies the normally dominant facilitatory effect. Kraut and Smothergill replicated a prior finding that prefamiliarization with one of two colored disks slows down responses to the prefamiliarized disk on subsequent test trials. However, in accord with their two-factor theory, this effect was reversed when the interval between the prefamiliarization and test trials was increased from 0 to 30 min, allowing the inhibitory factor to dissipate and the opposing facilitatory factor to dominate. In a follow-up study, Kraut, Smothergill, and Farkas (1981) observed comparable effects with word stimuli. Prefamiliarization retarded reaction time to words under conditions designed to more heavily weight the inhibitory factor but enhanced reaction time under conditions designed to more heavily weight the facilitatory factor.

Other evidence that bottom-up processing is suppressed for prefamiliarized stimuli was reported by Lorch, Anderson, and Wells (1984). They found that the speeded categorization of target shapes was enhanced if the distractor shapes with which they were paired had appeared repeatedly in a prefamiliarization phase. These findings support the suggestion of Cowan (1988) that habituation of distractor stimuli serves as a filter in selective attention tasks. Finally, an effect suggestive of perceptual habituation is that of *latent inhibition*, in which prefamiliarization decreases the conditionability of stimuli on subsequent classical conditioning trials (see, e.g., Lubow, 1973; see Hall, 1991, for an account of latent inhibition that is similar to mismatch theory).

Semantic satiation. A phenomenon related to perceptual habituation of individual stimuli is the *semantic satiation* of whole conceptual categories. For example, Smith (1984) had subjects repeat a category name (e.g., *fruit*) either 3 or 30 times prior to deciding whether or not a test word is an instance of the category. Repetition of the category name yielded a slowdown in decision times for positive instances (e.g., *apple*) but either no change or even a speedup for negative instances (e.g., *robin*). Smith and Klein (1990) replicated these effects and extended them to a category matching task. For example, repetition of *fruit* slowed down the matching of *apple-pear* but had no effect on *robin-sparrow*. In addition, category repetition yielded a decrease in the usual facilitative effect of congruent flankers (e.g., *pear*) on decision time for individually presented instances (e.g., *apple*). However, decision time for both single and paired instances of prefamiliarized categories continued to remain shorter than that for control pairs even after 30 repetitions of the category name. Thus, in general accord with the idea that expectation has opposing effects on perceptual processing, the latter, positive effect of conceptual priming outweighed the negative effect of semantic satiation.

Memory inhibition. Expectancy-based perceptual inhibition should have memorial consequences. The spacing effect on explicit memory may be a case in point. Memory for repeated items in a long list is retarded to the extent that the repetitions are massed together rather

than distributed throughout the list (see, e.g., Hintzman, 1974). Johnston and Uhl (1976) provided evidence that the spacing effect is due, at least in part, to a perceptual decrement for massed repetitions. These investigators measured the amount of attention paid to repeated items on the basis of simple reaction time to subsidiary signals. Attention decreased sharply across four massed presentations of the same word but increased slightly across four distributed presentations.

Inhibition of return. Tipper, Brehaut, and Driver (1990) presented a target object in one of four display locations and a distractor object in one of the remaining locations. Time to localize the target on one trial was retarded if it appeared in the location that the distractor had occupied on the immediately preceding trial. Shapiro and Loughlin (1993) replicated this negative-priming effect and extended it by showing a degree of retardation even when the target reappeared in the same location that it had occupied on the previous trial.

This latter finding is very similar to the *inhibition of return* effect originally introduced by Posner and Cohen (1984) and replicated and extended by other investigators (see, e.g., Maylor & Hockey, 1985; but see Klein, 1993, for boundary conditions). Posner and Cohen examined simple reaction time to the onsets of simple peripheral signals. Provided that the intersignal interval was long enough (e.g., 300 msec) for attention to be fixed on a neutral location between signals, reaction time to the second of two successive signals was slower when it occurred in the same location as the first signal than when it occurred in a different location. This inhibition was associated with environmental rather than retinal loci. Posner and Cohen suggested that inhibition of return may function adaptively to optimize visual foraging. Immediately successive orientations to the same environmental location would put organisms at a disadvantage in the search for prey and in monitoring for predators.

Repetition blindness. The literature reviewed above indicates that repetition can suppress bottom-up processing for both the identities and locations of objects. Can this suppression be so strong as to cause people to fail to see the repeated object altogether? The phenomenon of *repetition blindness* suggests that the answer is "yes." For example, Kanwisher and her colleagues have shown that the second occurrence of a word or object repeated in a rapid visual sequence of other words or objects often fails to be subjectively experienced (see, e.g., Kanwisher, 1987; Kanwisher & Potter, 1989, 1990). This can occur even if the omission of the repetition turns a grammatical sentence into a nongrammatical one. For example, a sentence such as "When she spilled the ink there was ink all over" tended to be reported as "When she spilled the ink there was all over." The effect tends to occur for spatial as well as for sequential arrays of items, when there is a case change between the two occurrences of an item, and even when the second item is a homograph (e.g., "He wound the cloth around the wound in his hand") or homonym (e.g., "She rose from the chair to take the red rose

from the bouquet") of the first (see, e.g., Humphreys, Besner, & Quinlan, 1988; Kanwisher, 1991, Kanwisher & Potter, 1990).

Primed inhibition. Semantic priming is defined by the positive effect that a prime word such as *bread* has on the response (e.g., pronunciation or lexicality judgment) to a semantically related test word such as *butter*. This phenomenon has become one of the most popular means of assessing the automatic spreading activation of associated mental representations that perceptual processing is assumed to entail (see, e.g., Neely, 1991; Schvaneveldt & Meyer, 1973). On the other hand, there are conditions under which primes seem to inhibit, rather than activate, related representations. For example, in their research on the tip-of-the-tongue phenomenon, Brown and McNeill (1966) found that definitions can sometimes impede word recall. Similarly, Slamecka (1968) found that the provision of a random subset of words from a list can impede recall of the rest of the words. An inhibitory effect has been observed by Dagenbach and his colleagues for both masked and poorly learned primes (Dagenbach, Carr, & Barnhardt, 1990; Dagenbach, Carr, & Wilhelmsen, 1989) and by Lambert and Voot (1993) for primes flanking the test word in the right hemisphere. Collectively, these findings indicate that an inhibitory effect of expectancy can surface when the degree and precision of expectancy are low.

Even when priming is facilitatory in terms of semantic or conceptually driven processing, the findings of Besner, Smith, and MacLeod (1990) suggest that it may be inhibitory in terms of physical or data-driven processing. These investigators observed the usual positive effect of primes on lexical decisions for related test words, but they observed a negative effect on time to decide whether the prime and test words had letters in common. Likewise, whereas explicit memory for semantically primed words is enhanced, implicit memory is sometimes inhibited (Jacoby, 1983). Thus, semantic priming tends to have a positive effect on measures more reflective of the amount of conceptually driven processing, such as lexical decision time and explicit memory, and a negative effect on measures more reflective of the amount of data-driven processing, such as letter matching and implicit memory (see Farah, 1989).

Word-inferiority effect. Interactive-activation theorists often cite the word-superiority effect as evidence for enhanced data-driven processing of expected inputs. This effect is defined by the observation that a single letter, such as *d*, is sometimes perceived more readily when it is presented in the context of a word, such as *word*, than when it is presented in the context of a nonword, such as *rowd*, or even by itself (see, e.g., Reicher, 1969). We suggest that this effect arises under conditions in which the enhanced conceptually driven processing of contextually constrained letters outweighs their suppressed data-driven processing.

That the net effect of contextual constraint can sometimes be inhibitory is indicated by the findings of Healy

and her colleagues in which the detection of a letter, such as *t*, in a passage of text is often lower when it is embedded in a very familiar word context, such as *the*, than in a less familiar word context, such as *tar* (see, e.g., Healy, 1976; Healy & Drewnowski, 1983). Drewnowski and Healy (1977) attribute this word-*inferiority* effect to the premature cessation of letter processing in unitized words. In line with mismatch theory, this effect has been found to diminish to the extent that conceptually driven processing of a word is disrupted by such procedures as presenting words in a list format, requiring a right-to-left scan direction, presenting a long blank space before or after the word, or presenting the word in isolation (see, e.g., Hadley & Healy, 1991).

Other evidence for the suppressed bottom-up processing of contextually constrained words comes from studies of parafoveal processing of words during text processing. For example, O'Regan (1979) showed that the word *the* is often identified parafoveally and skipped over by foveal fixations. Similarly, Ehrlich and Rayner (1981) found that words are more often skipped in high-constraint passages than in low-constraint passages. In brief, whereas the identification of individual letters may be enhanced by high contextual constraints on *letter* processing (e.g., relatively unfamiliar and isolated words), it may be inhibited by high contextual constraints on *word* processing (e.g., very familiar words in textual passages).

Registration without learning. If conceptually driven processing is enhanced but data-driven processing is inhibited for familiar objects, observers may be facile at identifying such objects but not at describing their physical details. In a classic study, Nickerson and Adams (1979) showed that people have little detailed knowledge of the physical appearance of common objects that they may have experienced and used thousands of times. In particular, people were found to have surprisingly poor memory for the appearance of pennies on a variety of tests. Even when they were supplied with the eight basic features that appear on the two sides of a penny, not 1 of 20 subjects was able to correctly configure these features. We suspect that people have poor specific memory for many other familiar objects, scenes, and events, such as the arrangement of letters on a telephone dial and the words to popular songs, such as national anthems.

Hintzman, Curran, and Oppy (1992) presented subjects long lists of pictures or words, in which some items were presented up to 25 times. Although judged frequency of presentation increased monotonically with actual frequency, the discriminability of a repeated item from a similar nonrepeated item remained constant. Hintzman et al. attributed these findings to *registration without learning*:

On repetitions, subjects apparently acquired information useful for the frequency judgment task without acquiring any more of the detail required for a fine-grained discrimination. It is as though if they did not encode such detailed information on the first presentation, they did not encode it at all . . . we suspect that we have stumbled upon evidence of a characteristic way that humans process infor-

mation in everyday life. To get around in the environment, people do not generally need to analyze the structure of objects once they have become familiar . . . indeed, further processing of such stimuli might be wasteful of valuable cognitive resources—something that evolution may have programmed humans to avoid. (p. 679)

Expertise-based inhibition. Although the facilitative effects of expertise are well-known (see, e.g., Chase & Simon, 1973; Voss & Bisanz, 1985), it is possible that these effects, like those of stimulus repetition and priming, are due to conceptually driven processing and tend to conceal a counteracting inhibitory effect of expertise on data-driven processing. In line with this possibility, Myles-Worsley, Johnston, and Simons (1988) found that recognition memory for normal chest X rays first increased and then declined with increasing radiographic expertise. Schmidt and Boshuizen (1993) observed a similar curvilinear effect in a study of the relationship between medical knowledge and memory for clinical case reports. However, this effect was eliminated when bottom-up processing was curtailed by a reduction in the time given to study the case reports. When study time did not permit a full reading of the reports, memory was directly related to medical expertise. Interestingly, although recall increased with study time for nonexperts, it remained constant for experts. Apparently, because the experts were not heavily dependent on bottom-up processing in their study of the case reports, their performance was not heavily dependent on the opportunity to engage in such processing.

We suspect that inhibition of data-driven processing operates also in chess and other previously studied domains of expertise. For example, after a quick glance at a chessboard during an actual game, a chess master may be able to reconstruct the configuration of pieces but fail to recall some of their incidental physical features.

Schema-based inhibition. As in the case of expertise, the usual facilitatory effects of active schemata on the processing of natural scenes and scripts may obscure an underlying inhibitory effect on bottom-up processing. One hint of this inhibitory effect is that whereas schematic perception enhances recall and recognition, it also yields a relatively high rate of importations and false alarms (see, e.g., Bower, Black, & Turner, 1979; Lewicki, Hill, & Sasaki, 1989). It is as though what people see in familiar environments can be determined as much or more by their expectations as by the actual data. However, because the bottom-up details of familiar contexts usually conform to expectation, the benefits of schematic perception are likely to outweigh any costs in terms of inhibition of bottom-up processing.

Evidence for the suppressed bottom-up processing of schema-consistent information has been generated by studies of the so-called *Moses illusion* and other failures of *anomaly detection* (see, e.g., Barton & Sanford, 1993; Erickson & Mattson, 1981). The Moses illusion refers to the frequent failure of people to notice the error in the question, "How many animals of each kind did Moses

put on the Ark?" Barton and Sanford showed that this sort of illusion increases with the extent to which the anomaly is consistent with the general theme or schema activated by the context. For example, they observed that detection of the anomaly in "The authorities were trying to decide where to bury the survivors" was substantially less when this sentence was read in the context of a report of an airplane crash than when it was read in the context of a bicycle accident. Barton and Sanford conjecture that when "the global fit of concepts in working memory . . . is high, then more detailed, effortful, time-consuming analysis may not take place" (p. 478).

Perhaps the most direct evidence for schema-based inhibition of bottom-up processing comes from the recent work of von Hippel, Jonides, Hilton, and Sowmya (1993). In line with the prevailing theme of this section, these investigators suggested "that although schemata facilitate organized conceptual processing, and hence recall, they simultaneously inhibit perceptual encoding. This inhibitory effect should emerge because schemata allow perceivers to rely on prior knowledge in place of incoming information" (p. 921). Capitalizing on prior evidence that bottom-up processing can be assessed in terms of implicit, perceptual memory (see, e.g., Jacoby, 1983), von Hippel et al. tested and confirmed the hypothesis that schematic processing of input material reduces perceptual memory of that material. For example, measures of perceptual identification, stem completion, and word-fragment completion for words from Bransford and Johnson's (1973) ambiguous "washing clothes" passage were lower for schematic readers (those given the disambiguating title) than for aschematic readers. This detrimental effect of schematic perception on implicit, perceptual memory of ambiguous passages stands in marked contrast to its typical facilitative effect on the explicit memory of these passages.

Summary. Various lines of evidence converge on the conclusion that conceptually driven facilitation of expected inputs is accompanied by data-driven inhibition. As conceptual nodes or representations are turned up, physical nodes or representations appear to be turned down. To the extent that a dependent variable depends on conceptual processing, the net effect of these opposing processes is likely to be facilitation. To the extent that a dependent variable depends on the bottom-up processing of physical details, the net effect is likely to be inhibition. Thus, prime words are likely to facilitate identification of test words but inhibit identification of their component letters, a preceding passage of text is likely to facilitate detection of the word *the* but inhibit detection of the letter *t*, repetition of pennies or words is likely to facilitate familiarity judgments but inhibit processing of their details, and active schemata are likely to facilitate apprehension of the gist of passages but inhibit memory for their details.

Perceptual Facilitation of Unexpected Inputs

We review next evidence that conditions associated with inhibited bottom-up processing of expected inputs are as-

sociated also with the enhanced bottom-up processing of unexpected inputs. This evidence has been generated by studies of the orienting response, mismatch negativity, memory facilitation, expertise-based facilitation, schema-based facilitation, and perceptual popout.

Orienting response. A rich empirical literature, much too massive to review in detail here, has accumulated in support of Sokolov's concept of a dynamic interplay between habituation and orientation (see, e.g., Berlyne, 1960; Sharpless & Jasper, 1956; Sokolov, 1963). The orienting response has often been used to measure attention in nonverbal organisms such as human infants and nonhuman animals (see, e.g., Cambell, Hayne, & Richardson, 1992). In general, this response is discriminable from startle responses and is characterized by an orientation of the receptors toward particular new or changed sources of stimulation. The response habituates with repetition of the stimulation but reappears upon the introduction of other changes. Thus, the orienting response betrays a special sensitivity of organisms to unexpected inputs.

Mismatch negativity. Various components of event-related potentials (e.g., N1 and P3) appear to index sensory or perceptual habituation in that they decline in magnitude across successive presentations of the same stimulus. Näätänen (e.g., 1990, 1992) has investigated a derived component of the event-related potential, called *mismatch negativity*, that is elicited only by a change in monotonous, or otherwise predictable, auditory stimulation. Borrowing from Sokolov (1963), Näätänen suggests that the adaptation series of tones establishes a neural model or expectancy against which subsequent tones are compared. Bottom-up processing is suppressed for tones that match the model but is accentuated for mismatching tones. The detection of mismatches occurs very early and is revealed in terms of mismatch negativity. The fact that mismatch negativity is not elicited just by new stimulation, such as the first tone in a series, indicates that it is not reflective of change per se. The fact that it occurs even when the same tone occurs in succession during a sequence of alternating tones suggests that it is not based on simple neural refractoriness. Rather, mismatch negativity indicates a special sensitivity of the mind/brain system to departures from expectation.

Memory facilitation. In their diagnosis of the Brown-Peterson paradigm, Keppel and Underwood (1962) demonstrated that the characteristic rapid forgetting of single items when rehearsal is prevented is attributable to a buildup over trials in proactive inhibition. Subsequent work, especially by Wickens (e.g., 1970), demonstrated that proactive inhibition continues to build up only if the succession of items are drawn from the same "class," such as animal names. Thus, the buildup of proactive inhibition may be another example of the memory- and primed-inhibition phenomena noted above. However, when a succession of several items from one class, such as animal names, is followed by an item from another class, such as vegetable names, then the new item is *released* from proactive inhibition. Although the buildup and release

from proactive inhibition may be a consequence of several factors, including trace discriminability and accessibility (see, e.g., Gardiner, Craik, & Birtwistle, 1972), one of the factors may be the inhibition and disinhibition of data-driven processing during item encoding.

Another memory phenomenon to which perceptual inhibition and disinhibition may contribute is the well-known von Restorff effect in which free recall is enhanced for physically or semantically unique or isolated items—for example, a word printed in blue ink in a list of words printed in red ink or an animal name in a list of vegetable names (for a review, see Wallace, 1965). In general, the size of the effect is a function of the degree of uniqueness. For example, if the list contains two words printed in blue ink, the size of the effect is greater for the first word. Karis, Fabiani, and Donchin (1984) measured the amount of attention paid to words in terms of the P3 component of the event-related potential. This measure was greater for isolates, in terms of type size, than for nonisolates. Thus, the von Restorff effect is at least partially attributable to perceptual, as opposed to strictly memorial, processing.

Expertise-based facilitation. In the Myles-Worsley et al. (1988) study of radiological expertise, the most experienced radiologists showed not only exceptionally poor memory for normal radiographs but exceptionally good memory for abnormal radiographs (those indicating some sort of pathology). Thus, those radiologists who were the most familiar with the appearance of normal chest X rays demonstrated both the most memory inhibition of the normal images and the most enhancement of the deviant images. Radiographic expertise apparently was not related to general memory ability, since a uniformly high level of human face recognition was attained by all groups.

Because nearly everyone is an expert at processing faces, face memory may also afford evidence of sensitivity to unexpected inputs. Indeed, the available evidence indicates that this everyday form of expertise involves the same processing dynamics as do more specialized forms (see, e.g., Diamond & Carey, 1986). For example, high levels of face memory require at least 10 years of experience with faces, and atypical faces are better remembered than typical ones. If the superior memory for atypical faces is like the expert radiologists' superior memory for abnormal radiographs, it should arise from both the suppressed processing of expected facial features and the enhanced processing of unexpected features.

Support for this hypothesis was reported by Locksley, Stangor, Hepburn, Groszovsky, and Hochstrasser (1984, Experiment 3). In a study phase, subjects viewed slides of faces, some of which had been prejudged to typify introverts and the rest of which had been prejudged to typify extroverts. The distinctions between the introverted and extroverted faces were based most of all on the mouth and eyes, called *schema-relevant* features, and least of all on the nose and chin, called *schema-irrelevant* features. Subjects were led to expect one or the other type of face by having that type represented in 75% of the study slides.

In a subsequent recognition memory test, subjects were shown old and new intact faces, faces containing only schema-relevant features, and faces containing only schema-irrelevant features. In agreement with prior effects of face typicality, memory was better for unexpected faces than for expected faces and for intact faces than for partial faces. It is important to note that memory for partial faces depended on whether they contained expected or unexpected features. Memory was better for the schema-relevant features of expected faces but was better for schema-relevant features than it was for schema-irrelevant features of unexpected faces. This interaction is consistent with the following interpretation: In the study phase, subjects focused initially on the schema-relevant features (i.e., mouth and eyes) of each face. Further bottom-up processing of these features was inhibited when they agreed but was enhanced when they disagreed, with expectation.

Schema-based facilitation. The sensitivity of experts to unexpected input is in accord with the much larger literature on schematic processing of natural scenes and temporal events. It is now well-established that attention and memory in natural contexts are greater for unexpected or incongruous objects and events than for expected or congruous ones (see, e.g., Bower et al., 1979; Friedman, 1979; Stangor & McMillan, 1992). For example, Bower et al. showed that memory for unexpected events in short stories of everyday activities (e.g., a visit to a dentist) is superior to that for expected events. Similarly, both Friedman (1979) and Loftus and Mackworth (1978) showed that when people scan drawings or pictures of common scenes (e.g., barnyards), they tend to fix their eyes longer and more often on incongruent objects (e.g., an octopus) than on congruent objects (e.g., a tractor). These effects appear to be based more on processing facilitation for unexpected stimuli than on inhibition for expected ones, because measures of memory and attention tend to be higher for both expected and unexpected stimuli than for neutral or schema-irrelevant ones (e.g., a tree in a barnyard scene).

Perceptual popout. Is the processing facilitation of unexpected inputs attributable to the enhanced detectability of these inputs in the first place or to the special, controlled processing accorded them after they are detected? Some of the effects reviewed above are at least partially attributable to the latter. For example, Srull (1981) showed that the enhanced memory for schema discrepancies disappeared when the opportunity for controlled processing during the study phase was delimited by requiring subjects to perform a subsidiary task. Similarly, Biederman, Mezzanotte, and Rabinowitz (1982, Experiment 1) limited viewers to only 150-msec glimpses of naturalistic scenes (e.g., a kitchen) and found accuracy of localization to be substantially less for unexpected objects (e.g., a fire hydrant) than for expected objects (e.g., a toaster). Following an interactive-activation line of thinking, Biederman et al. suggested that a schema-driven

enhancement of perception of the expected objects yielded a perceptual inhibition, or sink-in, of the unexpected objects.

On the other hand, various perceptual popout phenomena indicate a genuine perceptual sensitivity to discrepant components of visual arrays. For example, in a search for the target letter *d* in an array of letters that contains a "singleton," or discrepant letter, with respect to an irrelevant property such as sudden onset or color, detection is often enhanced if the target is the singleton but inhibited if a distractor is the singleton (see, e.g., Folk, Remington, & J. C. Johnston, 1993; Yantis, 1993). Although these effects might be attributable to figure-ground contrast along dimensions other than expectancy or familiarity, novel popout cannot. Novel popout indicates that unexpected singletons in a familiar scene receive a bottom-up boost in perceptual processing. Our own research on novel popout is reinforced by earlier findings that an unfamiliar target form (e.g., upside-down letter) in a background of familiar distractor forms (e.g., rightside-up letters) is easier to detect than a familiar target form in a background of unfamiliar distractor forms (Frith, 1974; Reicher, Snyder, & Richards, 1976). If Biederman et al. (1982) had used an all-unexpected control condition, they might have been able to detect a novel popout effect. For example, a fire hydrant might be more readily perceived in the context of a kitchen than in the context of a random assemblage of objects (see Hawley et al., 1994).

Summary. The same conditions associated with the data-driven inhibition of expected inputs appear to be associated also with the data-driven facilitation of unexpected inputs. Thus, perceptual habituation and satiation of repeated stimuli or stimulus classes yield perceptual and memorial facilitation of new or changed stimuli or stimulus classes. Similarly, familiar scenes, scripts, and other inputs (e.g., human faces) are associated not only with suppressed data-driven processing of expected components but also with perceptual and memorial popout of deviant components.

PART 4 A Computational Instantiation of Mismatch Theory

The extant empirical literature affords converging evidence in support of the assumptions of mismatch theory that data-driven processing is both inhibited for expected inputs and enhanced for unexpected inputs. We now describe in more detail how these assumptions are computationally realized, and we show how the entire pattern of novel-popout effects emerges from simulation runs of the model.

Model Architecture and Processing Dynamics

Some of the computational dynamics of mismatch theory in the context of our novel-popout paradigm are il-

lustrated in Figures 2-5. The levels of node activation shown in these figures were derived directly from the simulation runs of the model through the all-novel (Figures 2 and 3), all-familiar (Figure 4), and one-novel (Figure 5) arrays of Experiment 4 of Johnston et al. (1990). As Figure 2 shows, bottom-up processing of an all-novel array proceeds through two tiers of processing units or nodes. The first tier consists of *iconic nodes* arranged in a lateral inhibitory network. This iconic tier extracts in parallel all the information needed to identify and localize objects and relays the information to the second tier. Location information is sent to an assemblage of *location nodes*, and identity information is sent to an assemblage of *identity nodes*. The inclusion of both location and identity nodes in different modules in the upper tier is consistent with evidence that such modules of specialized neurons exist in the prefrontal cortex of real brains (see, e.g., Wilson, Ó Scalaidhe, & Goldman-Rakic, 1993). The model provides a different identity node for each novel and familiar object. Although the model is equipped with more than four location nodes, only four are needed for the simulations of our research to date. Yoked to the bottom-up activation links between the two tiers are top-down inhibition links. To the extent that a node on the upper tier is activated, inhibition is reflected downward to the iconic node with which it is connected. Hence, the activity of each iconic node is modulated from three directions: data-driven activation from the external array, lateral inhibition from other active iconic nodes, and top-down inhibition from active identity and location nodes.

The model starts off as a hardwired network consisting of three kinds of link: lateral inhibitory links within the iconic tier and both bottom-up excitatory and top-down inhibitory links between tiers. However, as bottom-up excitation continues to be generated from the iconic tier, a softwired network begins to evolve within the upper tier. As Figure 3 illustrates, two new kinds of excitatory links begin to form: one between the location and identity nodes coactivated by the same iconic node and, therefore, representing the same object, and the other between any two coactive identity nodes. The strength, or connection weight, of a soft link is a function of the level of coactivity of the two nodes. As this softwired network evolves across all-familiar trials, the upper-tier nodes begin to receive new sources of activation. In addition to bottom-up excitation from the iconic tier, the location nodes begin to receive excitation from the identity nodes with which they are linked, and the identity nodes begin to receive excitation from both the location and other identity nodes with which they are linked.²

A third kind of link, not shown in Figure 3, is a collateral link, or metalink. These form between coactive identity-location links. Metalinks subserve configurational learning such that the occurrence of one object in one location can prime the links that represent other objects in other locations. These links accelerate the conceptually

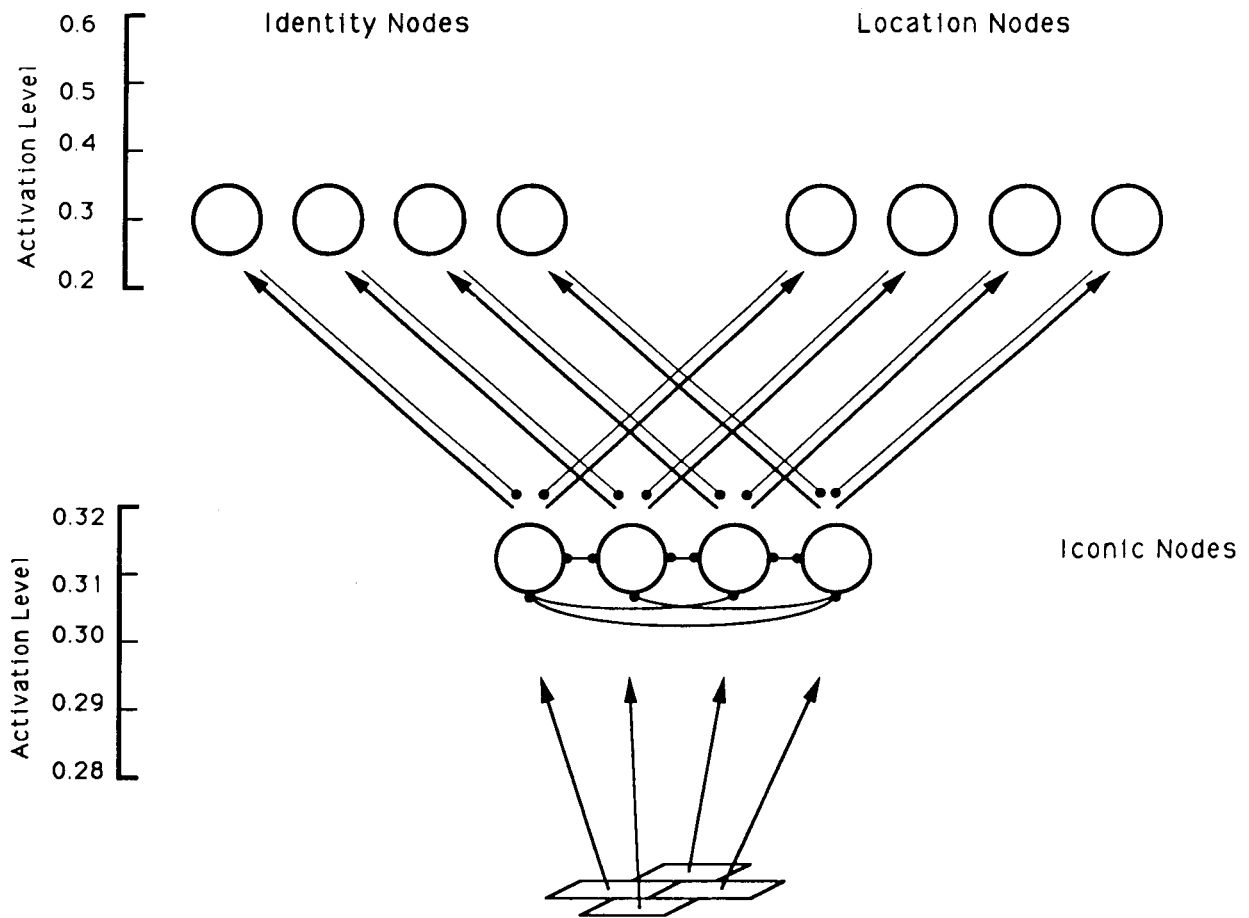


Figure 2. General architecture of the model. Excitation, indicated by lines terminating in arrows, flows upward from an array of four objects through iconic nodes to both location and identity nodes. Inhibition, indicated by lines terminating in closed circles, flows both laterally between iconic nodes and downward to iconic nodes from identity and location nodes. The levels of node activation were derived from simulation data for all-novel arrays. The relative thickness of a line represents the relative strength of a connection.

driven processing of familiar scenes that are spatially stable in comparison with those that are not, such as the all-familiar arrays in most of our studies.

Application of Model to Different Types of Array

The weak softwired network formed for an all-novel array soon decays. However, that formed for the first presentation of an all-familiar array grows stronger across array repetitions. The development of a strong softwired network across repetitions of an all-familiar array models the molding or attunement of the mind/brain system to the outside world. Thus, the propagation of excitation and inhibition across the composite soft- and hardwired network begins to differ for all-novel and all-familiar arrays. In comparison with all-novel arrays, the all-familiar arrays engender more spreading excitation of identity and location nodes. The identity nodes receive the most spreading excitation because they are linked with one another as well as with the location nodes. The higher excitation of the upper-tier nodes ricochets more inhibition

downward to the iconic nodes, and this, in turn, attenuates both the lateral-inhibition and bottom-up excitation emanating from these nodes. However, the spreading excitation of location and identity nodes across the learned links more than offsets their self-generated inhibition, yielding the standard baseline superiority of all-familiar over all-novel arrays. Thus, as a comparison of Figures 2-4 shows, the model remains relatively data-driven on all-novel trials but becomes progressively more conceptually driven on all-familiar trials.

The stage is now set to show how the conceptually driven suppression of data-driven processing can yield novel popout. A frozen cross-section of the model's dynamics for a one-novel array is depicted in Figure 5. The inhibition of the iconic nodes for the familiar field objects reduces the lateral inhibition that they convey to the iconic node for the lone novel object. This disinhibition has two consequences. First, it increases the coactivity of the location and identity nodes representing the novel object, strengthening the link between them and promoting novel popout. Second, it delivers more lateral inhibi-

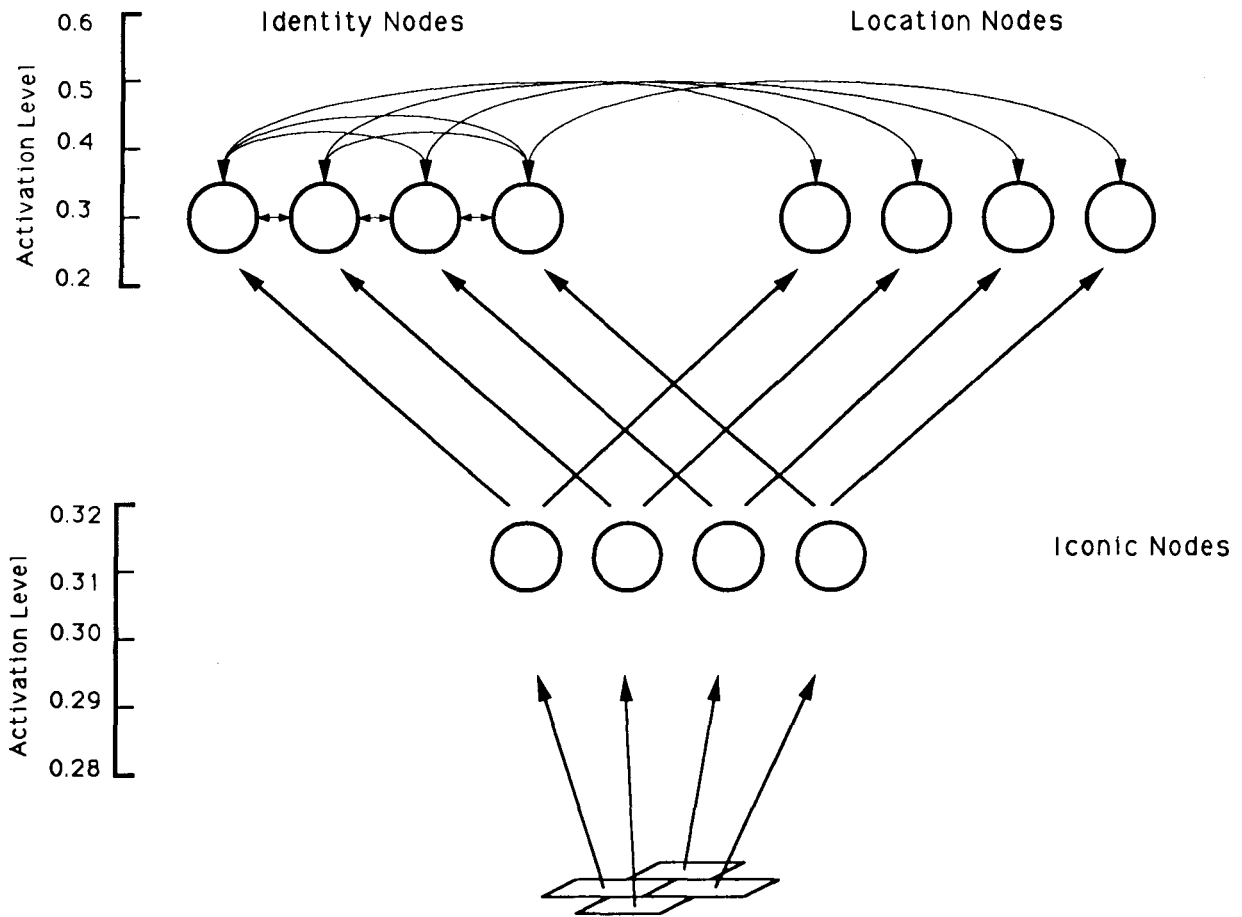


Figure 3. Formation of weak links between coactive identity nodes and between identity and location nodes activated by the same object from an all-novel array. For clarity, the inhibitory links shown in Figure 2 have been omitted.

tion to surrounding iconic nodes, further suppressing their activity and promoting familiar sink-in.

Response Production

Following the brief exposure of an array of words, the model is probed in the same way as are human observers. It is presented one of the words and must make a localization response. The data-driven processing of a probe launches a reverberatory spread of excitation of both identity and location nodes. The model makes a localization response on the basis of the relative activation levels of the four location nodes as they continue to fluctuate throughout the probe interval. When the difference in activation between the most active location node and the next-most active one reaches criterion, or when a maximum time interval elapses, the model's response is dictated by the most active node. This response-production algorithm is a variant of that developed by Cohen, Dunbar, and McClelland (1990).

Simulation of Novel Popout

The mathematical details and algorithms of the computational version of mismatch theory are presented else-

where (Hawley, Johnston, & Farnham, 1993). Suffice it to say that we set the values of the various learning, activation, decay, and response-production parameters on the basis of their ability to simulate the baseline effect on a short sequence of all-novel and all-familiar trials. Using these parameter values, we then ran the model through Experiment 4 of Johnston et al. (1990). A comparison of the simulation data, summarized in the right panel of Figure 1, with the empirical data, summarized in the left panel, reveals a reasonable match. The complete pattern of effects emerged, including the accuracy difference between novel and field words in the one-novel arrays. The match no doubt would be even closer if we had fitted our parameters to the actual data of Experiment 4. A more detailed examination of the model's performance revealed further evidence of its simulative power. Two independent versions of the original experiment were constructed, and each was administered to half of the subjects. Although the same pattern of results obtained for both versions, it was more conspicuous in one than in the other, indicating the occurrence of trial sequencing effects. The model displayed a comparable difference between versions.

Curiously, the model's success is attributable in part to various subtle and emergent features of its dynamics

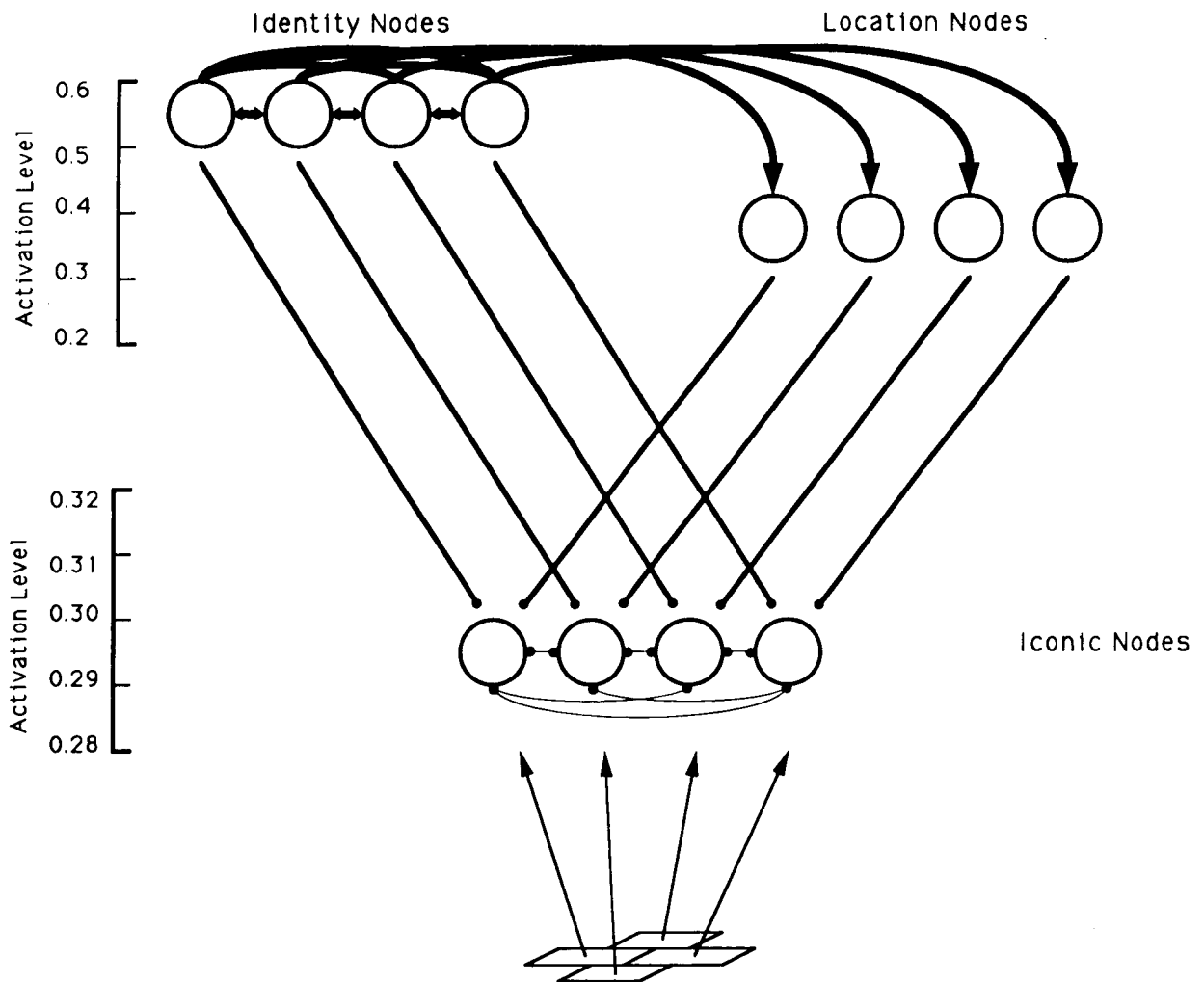


Figure 4. The model's response to all-familiar arrays. The formation of strong links across repetitions of these arrays leads to increased spreading activation between identity and location nodes and, in turn, to increased top-down inhibition of iconic nodes and suppressed lateral inhibition between iconic nodes. For clarity, both the bottom-up links and all but the most recently activated links between identity and location nodes have been omitted.

that we did not entirely anticipate. We note only two of these here. First, the model simulates the advantage of novel intrusions over field words in terms of localization accuracy despite the fact that, as Figure 5 reveals, the identity and location nodes representing the novel words are less active and more weakly interlinked than are those representing the field words. This apparent discrepancy arises from the fact, noted above, that the response of the model to a probe object is determined directly, not by the levels of node activation during the presentation of the one-novel array itself but rather by the ability of the probe to discriminate the correct location node from the incorrect ones. Thus, although the absolute connection weights between identity and location nodes are greater for the familiar field objects, the relative weight of the correct connection is, on average, greater for the lone novel objects.

Another unanticipated outcome of the model's dynamics is that the level of activation of the identity nodes for fa-

miliar objects is less in one-novel arrays than in all-familiar arrays despite the fact that the level of bottom-up activation of these nodes from the iconic nodes is somewhat greater in the one-novel arrays. Apparently, the absence of one of the familiar objects yields a reduction in the activation that reverberates across the three familiar objects that remain, and this reduction contributes to familiar sink-in.

Evaluation of Mismatch Theory

Although its underlying dynamics are more complicated than we had anticipated, mismatch theory simulates the basic pattern and certain details of novel-popout effects. We now consider how well mismatch theory compares with two popular alternatives, interactive-activation and resource-allocation models.

Interactive-activation theory. As already noted, interactive-activation models are diametrically opposed to mismatch theory. In these models, bottom-up process-

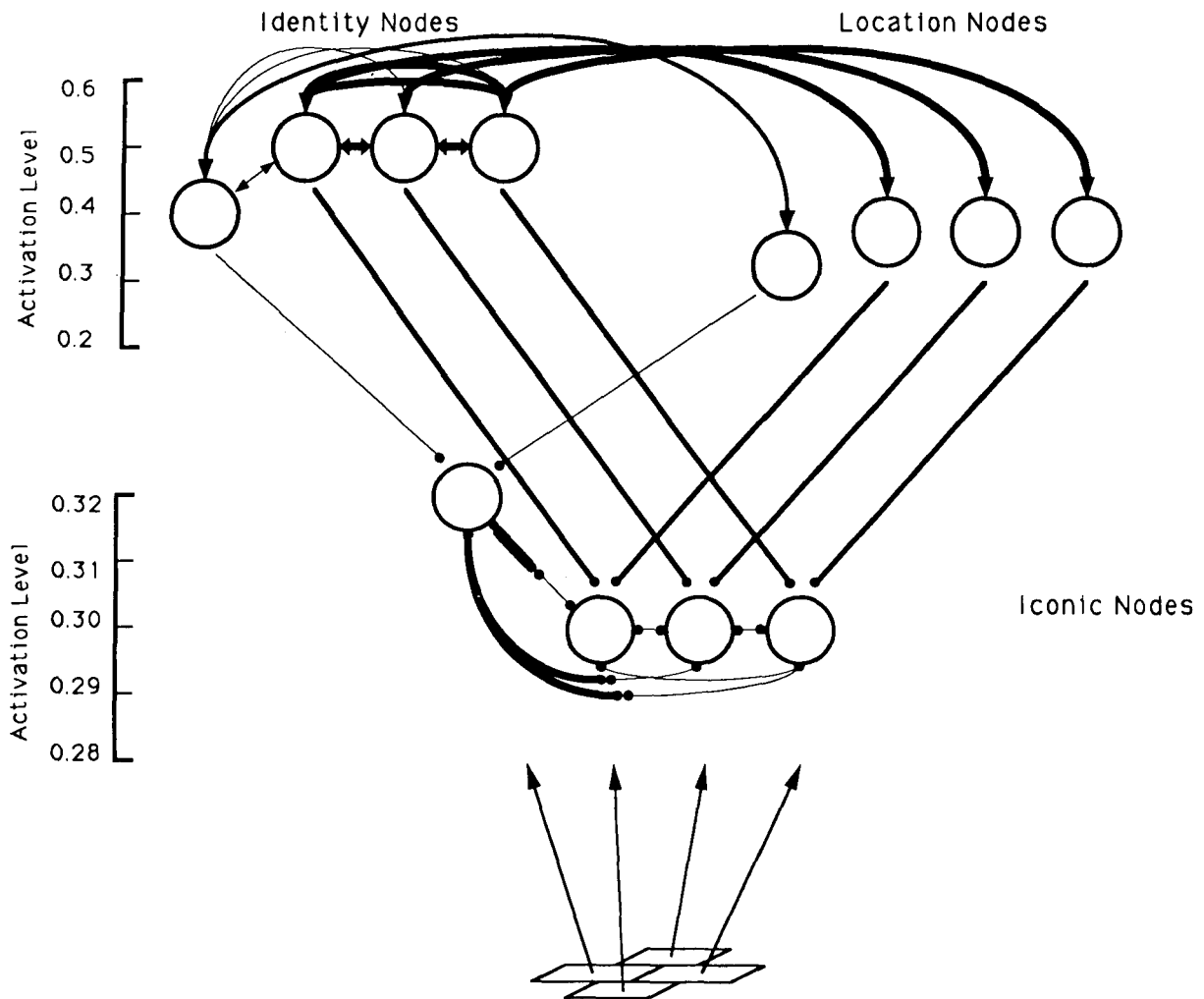


Figure 5. The model's response to one-novel arrays. The top-down inhibition of iconic nodes representing the familiar objects decreases the lateral inhibition converging on the iconic node representing the novel object. As a result, the latter node discharges more lateral inhibition to the former nodes and generates more bottom-up excitation of, and a stronger link between, the identity and location nodes representing the novel object. For clarity, both the bottom-up links and all but the most recently activated links between identity and location nodes have been omitted.

ing is enhanced for expected inputs and inhibited for unexpected inputs, a process that would produce novel sink-in rather than popout. Thus, interactive-activation theory cannot accommodate the main phenomenon that mismatch theory was designed to explain. On the other hand, mismatch theory can in principle readily accommodate the main phenomena that interactive-activation theory was designed to explain—namely, word-superiority and perceptual-restoration effects.³

Mismatch theory explains the word-superiority effect and perceptual restoration in terms of the enhanced conceptually driven processing of letters in word contexts. For example, owing to the ordinary links between regularly coactivated identity nodes, the context *wor* should prime the identity node for *d*. Moreover, owing to the formation of metalinks, the context *wor* should specifically prime the identity-location link that represents *d* in

the fourth location. Although all of this conceptually driven facilitation of the representation of *d* in the fourth location would discharge a proportional amount of top-down inhibition of the iconic node for *d*, the net effect could be one of perceptual facilitation, in which case both the word-superiority effect and perceptual restoration would result. That is, the perception of *d* in the fourth position could be facilitated by the context of *wor*, both when that position is occupied by *d* and when it is empty. Thus, mismatch theory appears to have the edge over interactive-activation theory in its ability to accommodate word-superiority and perceptual-restoration effects as well as novel-popout effects.

Resource-allocation models. Some of the reviewers of our presentations of mismatch theory, including the present one, have suggested that novel popout can be explained more simply in terms of the strategic allocation

of processing resources to novel and familiar inputs by a central executive or attention director. Because the perception of familiar inputs requires fewer resources than that of novel inputs, the executive can allocate to novel intrusions resources that are not needed for familiar field objects. In other words, the executive can withhold a portion of its attention from the field objects and focus it on any novel singletons. This is very much the same kind of explanation that has been offered to us in natural-language terms by some of our nonpsychologist acquaintances who assert that since one does not have to pay attention to familiar objects, one can focus it on unexpected objects. We submit that such natural-language interpretations of novel popout are circular and vacuous. They are circular because they essentially assert that novel objects popout (i.e., receive more attention) because they receive more attention (i.e., pop out). They are vacuous because they do not explicate the nature of the executive and its resources. What is attention and how does the executive so ingeniously manage it? How does the executive discriminate between novel and familiar inputs? These "explanations" encounter the same problem of infinite regress that characterizes all appeals to intelligent processing homunculi. The mystery of how the big mind works is not solved by equipping it with a little mind.

Resource-allocation models offer little more than an intuitive framework and vocabulary with which to describe mental phenomena. However, they fail to discriminate between opposing computational models and have no predictive power of their own. For example, these models would not be compromised if novel objects were, in accordance with interactive-activation models, observed to sink into, rather than pop out from, familiar fields. They would explain that the executive pays more attention to expected than to unexpected inputs. Indeed, to a limited extent, mismatch theory can be translated into resource-allocation terms. The inhibited bottom-up processing of expected inputs may be considered as a withdrawal of attention, and the accentuated bottom-up processing of unexpected inputs may be considered as a focusing of attention. Attention, though, is an emergent property of explicit computational dynamics; it is a phenomenon of the whole mind/brain system, not an unspecified resource managed by a mysterious little mind inside the big mind.

Summary of Mismatch Theory

Mismatch theory illustrates how conceptually driven inhibition of data-driven processing can serve both the stability and plasticity functions of mind/brain systems. The stability function is carried out by the overlaying of a softwired network of nodes over a hardwired network. As mental attunement proceeds and the softwired network becomes more active, there is a proportional shutdown of data-driven processing. This shutdown has at least three adaptive effects: It promotes mental stability (indexed by the baseline effect), it promotes mental plasticity (indexed by novel popout and familiar sink-in), and it accomplishes

both feats economically, in a way that obviates the need for special processing mechanisms and that minimizes the energy requirements of perceptual processing. All three effects are emergent by-products of the same underlying dynamics. The match between the simulation data and the empirical data suggests that the model and the human mind share many of the same dynamics, an important one of these being the suppressed bottom-up processing of expected inputs. Mismatch theory simulates a broader range of phenomena than do interactive-activation models and is more explicit computationally than are resource-allocation models.

PART 5 General Discussion

We began this paper by underscoring Treisman's (1992) observation of the mind's impressive ability to be molded by its past experiences without losing its sensitivity to new experiences. The mind/brain manages to be biased simultaneously toward both expected and unexpected inputs. We attributed these apparently opposing biases to the pivotal role that perceptual inhibition can play in maintaining a balance between mental stability and mental plasticity. Because the mind/brain can perceive familiar habitats in terms of conceptually driven processing, it can reserve data-driven processing for unexpected inputs. The balance between mental stability and mental plasticity allows organisms to carry out their routine activities in an energy-efficient way and still remain open to significant environmental changes. We consider below some of the implications of this view for the phenomenology of perception, for dynamic-systems theory, and for mental health and individual differences.

Phenomenology of Perception

The idea that perception of unperturbed, familiar environments is, for the most part, conceptually driven may lack intuitive appeal. For example, we can look around our offices and attend to any number of details. We can see the shelves of books and other assorted items, and they appear very much as we had expected. Certainly, such experiences are not just conceptually driven illusions. Our thesis is not that we cannot engage in substantial data-driven processing of our habitats but, rather, that we usually do not. In order to deal with the uncertainties and vagaries of our work, such as are entailed in the preparation of this paper, we rarely take explicit note of the details of our offices. Similarly, in walking across campus and digesting the sometimes unpredictable conversations of our colleagues, we seldom note the bottom-up details of the fountains, buildings, and passersby. What then, in these instances, is the nature of our subjective experience of the environmental contexts? It is probably not possible to answer this question experientially because the moment that one begins to analyze one's subjective experience of the context, the mix of conceptually driven

and data-driven processing is likely to suddenly shift to relatively more of the latter, thereby causing the subjective experience to change.

Although introspection and phenomenology may be of limited utility as methods of investigating the nature of ordinary perception in familiar environments, it is likely that most of us have had experiences that are consistent with the suppressed data-driven processing of expected inputs. For example, a common experience in driving along familiar routes is the sudden discovery that one has driven for several miles but has no explicit memory for what transpired. One of us was recently staring out the window of his office, deep in thought about this paper, only to suddenly realize that he had not "seen" a thing. On another recent occasion, he saw his cat in one of its favorite locations on a couch only to have this perception transformed a moment later into the stack of newspapers that actually lay there.

Do only novel intrusions and other object singletons, such as sudden onsets, capture our attention (i.e., data-driven processing)? Are we subjectively blind in familiar contexts that have no such intrusions or irregularities? We suspect that the answer is negative to both questions for at least two reasons. First, the evidence indicates that data-driven processing in familiar contexts is only turned down, not off. Second, the perceptual system may periodically revert to data-driven processing in order to verify and, when necessary, update expectancies. We had at one time entertained a third possible reason. Specifically, in most everyday situations, we may be primed by our transitory goals and tasks to attend to specific aspects of our familiar environments. Such naturalistic priming defines the *effect theory* of attention articulated by William James (see Johnston & Dark, 1986). Although there is some evidence for what may be called *primed popout* (e.g., Dark & Vochatzer, 1992), evidence to the contrary has been generated in our own laboratory by DeWitt (1994).

DeWitt's (1994) subjects performed our standard localization task on a long series of all-novel arrays. Each array was preceded by a prime word. Different types of array were defined by different relationships of the four words to the prime. All four words were associates of the prime in an *all-primed* array, none was in an *all-nonprimed* array, only one was in a *one-primed* array, and only one was not in a *one-nonprimed* array. The words in each array were semantically interrelated except for the one-primed and one-nonprimed singletons in the latter two types of array. Accuracy of localization was higher for all-primed arrays than for all-nonprimed arrays. This baseline effect is comparable to that found in our novel-popout studies (see Figure 1), except that it indicates a net facilitatory effect of short-term, as opposed to relatively long-term, expectancies. In addition, in the case of one-nonprimed arrays, performance rose above the all-nonprimed baseline for the nonprimed singletons but fell below the all-primed baseline for the primed field words. These effects are comparable to the standard novel-popout and familiar sink-in effects described in Part 1.

Surprisingly, these effects extended to the one-primed arrays. That is, performance rose above baseline for the nonprimed field words but fell below baseline for the lone primed words. Rather than indicating primed popout, these data describe primed sink-in, represent another instantiation of primed inhibition, and constitute strong additional evidence for expectancy-based inhibition of data-driven processing.

In brief, then, we propose the counterintuitive hypothesis that perceptual processing and the subjective experience thereof in familiar environments are more conceptually driven than data driven. We are not prepared to speculate about the exact mix of the two kinds of processing, but we suspect that the data-driven portion is considerably less than is commonly supposed.

The Mind as a Dynamic System

From the perspective of dynamic-systems theory, nature is not a static, Cartesian machine but is continually self-organizing and evolving, even as it slowly winds down (Jantsch, 1980; Prigogine & Stengers, 1984). For a system to prolong its life and vitality, it must strike a balance between self-perpetuation (e.g., stability) and transformation (e.g., plasticity). Self-perpetuation is indicated by the tendency of a system to settle into attractor states—states that are attuned to the environment and minimize the exportation of entropy (i.e., expenditure of energy). Because the environment of any one system is itself composed of other self-organizing systems, the continual, dynamic synchronization of the whole complex of systems ensures a degree of ecological balance and harmony. System transformation is indicated by the phase transitions, or bifurcations, that they undergo when they are perturbed sufficiently far out of their attractor states. Although a phase transition necessitates a transient increase in energy expenditure, it normally restabilizes the system and restores its relative quiescence.

Operating together, self-perpetuation and transformation allow for the evolution of complex systems—the *positive arrow of time*—even as the universe as a whole steadily succumbs to the second law of thermodynamics—the *negative arrow of time* (see, e.g., Coveney & Highfield, 1990). Because self-perpetuation and transformation have opposing effects, the integrity of the system depends on the maintenance of an essential tension between them. Too much stability can yield a closed, nonresilient system marooned in a deep attractor state; too much plasticity can yield an erratic, ephemeral system out of harmony with its environment. A prolonged imbalance in either direction eventuates in system deterioration.

As it is conceptualized in this paper, the mind is a dynamic system. The propensity of the mind to be conceptually driven and conform to familiar habitats represents mental self-perpetuation; its propensity to detect new or deviant features of these habitats promotes mental transformation. Like any other adaptive system, the mind manages to strike an optimal balance between these opposing processes in an energy-efficient manner. In the case

of the mind, the key to this balance appears to be the top-down suppression of expected inputs.

Mental Health and Individual Differences

From a dynamic-systems perspective, mental health is to be understood not in terms of functioning and malfunctioning parts of a mind/brain machine but rather in terms of the attunement of the mind/brain system to its environment. We suggest that a healthy mind/brain is one that maintains a degree of synchrony with a changing environment by striking an optimal balance between mental stability and mental plasticity. If the balance is tipped too far toward stability, the system could stagnate; if the balance is tipped too far toward plasticity, the system could disintegrate.

It is likely that even normal mind/brain systems occasionally swing too far in one direction or the other before the balance is restored. A swing too far toward stability might be indicated by one's proofreading ineptitude, especially with respect to one's own manuscripts, or by the object of one's search being in plain view but going unseen because it is not in its expected location. A swing too far toward plasticity might be indicated by absentmindedness and distractibility. Some empirically studied examples might include functional fixedness and problem-solving set for overstability (see, e.g., Levine, 1988) and action slips and mental lapses for overplasticity (see, e.g., Reason & Mycielska, 1982). Although transitory shifts too far toward either stability or plasticity are commonplace and usually innocuous, a chronic imbalance is likely to be maladaptive and the source of various cognitive and behavioral deficits and pathologies. Some possible instances of chronic overstability include the usual cognitive decrements associated with senescence and depression, and some possible instances of chronic overplasticity include attention-deficit disorders and schizophrenia.

The view of the mind/brain as a dynamic system has implications for intervention and treatment as well as for interpretation and diagnosis of cognitive and behavioral pathologies. Rather than trying to identify and repair the broken parts of a mind/brain machine, this systems view encourages a focus on, and possible manipulation of, the relationship between the mind/brain and its environment. An excessively stable mind/brain may need to be perturbed out of an obsolete or dysfunctional attractor state or prevented from sinking too far into it in the first place. An excessively plastic mind/brain may be in need of more environmental stability and predictability. Of the two sorts of imbalance, we suspect that excessive rigidity is the more prevalent one in contemporary Western society. Many individual human beings and the social and institutional systems in which they participate conspire to create stable, self-perpetuating attractor states. Our environments are engineered to be relatively stable and free of perturbation. Because of the relative blindness to bottom-up data that overstability yields, the mind/brain unwittingly joins the conspiracy and participates in its own potential demise. Thus, although perceptual inhibition of

expected inputs may be the key that can open closed minds, it may also be the key that can lock minds too tightly closed.

REFERENCES

- BARTON, S. B., & SANFORD, A. J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition*, *21*, 477-487.
- BERLYNE, D. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.
- BESNER, D., SMITH, M. C., & MACLEOD, C. M. (1990). Visual word recognition: A dissociation of lexical and semantic processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 862-869.
- BIEDERMAN, I., GLASS, A. L., & STACY, E. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology*, *9*, 22-27.
- BIEDERMAN, I., MEZZANOTTE, R. J., & RABINOWITZ, J. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*, 143-177.
- BOWER, G. H., BLACK, J. B., & TURNER, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177-220.
- BRANSFORD, J. D., & JOHNSON, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing* (pp. 383-438). San Diego: Academic Press.
- BROWN, R., & MCNEILL, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning & Verbal Behavior*, *5*, 325-337.
- CAMBELL, B. A., HAYNE, H., & RICHARDSON, R. (1992). *Attention and information processing in infants and adults: Perspectives from human and animal research*. Hillsdale, NJ: Erlbaum.
- CHASE, W. G., & SIMON, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.
- COHEN, J. D., DUNBAR, K., & MCCLELLAND, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332-361.
- COVENEY, P., & HIGHFIELD, R. (1990). *The arrow of time*. New York: Fawcett.
- COWAN, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163-191.
- DAGENBACH, D., CARR, T. H., & BARNHARDT, T. M. (1990). Inhibitory semantic priming of lexical decisions due to failure to retrieve weakly activated codes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 328-340.
- DAGENBACH, D., CARR, T. H., & WILHELMSSEN, A. (1989). Task-induced strategies and near-threshold priming: Conscious effects on unconscious perception. *Journal of Memory & Language*, *28*, 412-443.
- DARK, V. J., & VOCHATZER, K. G. (1992, November). Semantic priming can lead to selective attention. Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis.
- DEWITT, M. J. (1994). *Attention capture by primed and unprimed stimuli*. Unpublished doctoral dissertation, University of Utah, Salt Lake City.
- DIAMOND, R., & CAREY, S. (1986). Why faces are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107-117.
- DREWNOWSKI, A., & HEALY, A. F. (1977). Detection errors on the and: Evidence for reading units larger than the word. *Memory & Cognition*, *5*, 636-647.
- EHRlich, S. F., & RAYNER, K. (1981). Context effects on word perception and eye movements in reading. *Journal of Verbal Learning & Verbal Behavior*, *20*, 641-655.
- ERICKSON, T. D., & MATTSON, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning & Verbal Behavior*, *20*, 540-551.
- FARAH, M. J. (1989). Semantic and perceptual priming: How similar are the underlying mechanisms? *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 216-225.

- FOLK, C. L., REMINGTON, R. W., & JOHNSTON, J. C. (1993). Contingent attentional capture: A reply to Yantis (1993). *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 682-685.
- FRIEDMAN, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, **108**, 316-355.
- FRITH, U. (1974). A curious effect with reversed letters explained by a theory of schema. *Perception & Psychophysics*, **16**, 113-116.
- GARDINER, J. M., CRAIK, F. I. M., & BIRTWISTLE, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning & Verbal Behavior*, **11**, 778-783.
- GROSSBERG, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, **11**, 23-63.
- HADLEY, J. A., & HEALY, A. F. (1991). When are reading units larger than the letter? Refinement of the unitization reading model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1062-1073.
- HALL, G. (1991). *Perceptual and associative learning*. Oxford: Oxford University Press, Clarendon Press.
- HAWLEY, K. J., JOHNSTON, W. A., & FARNHAM, J. M. (1993, May). *Mismatch theory of novel popout: A computational model*. Paper presented at the Third West Coast Attention Conference, Eugene, OR.
- HAWLEY, K. J., JOHNSTON, W. A., & FARNHAM, J. M. (1994). Novel popout with nonsense strings: Effects of predictability of string length and spatial location. *Perception & Psychophysics*, **55**, 261-268.
- HEALY, A. F. (1976). Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 235-242.
- HEALY, A. F., & DREWNOWSKI, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 413-426.
- HINTZMAN, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories of cognitive psychology: The Loyola Symposium*. Potomac, MD: Erlbaum.
- HINTZMAN, D. L., CURRAN, T., & OPPY, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 667-680.
- HUMPHREYS, G. W., BESNER, D., & QUINLAN, P. T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General*, **117**, 51-67.
- JACOBY, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **3**, 300-324.
- JANTSCH, E. (1980). *The self-organizing universe*. New York: Pergamon.
- JOHNSTON, W. A., & DARK, V. J. (1986). Selective attention. *Annual Review of Psychology*, **37**, 43-75.
- JOHNSTON, W. A., HAWLEY, K. J., & FARNHAM, M. (1993). Novel popout: Empirical boundaries and tentative theory. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 140-153.
- JOHNSTON, W. A., HAWLEY, K. J., PLEWE, S. H., ELLIOTT, J. M. G., & DEWITT, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General*, **119**, 397-411.
- JOHNSTON, W. A., & UHL, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 153-160.
- KANWISHER, N. G. (1987). Repetition blindness: Type recognition without token individuation. *Cognition*, **27**, 117-143.
- KANWISHER, N. G. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 404-421.
- KANWISHER, N. G., & POTTER, M. C. (1989). Repetition blindness: The effects of stimulus modality and spatial displacement. *Memory & Cognition*, **17**, 117-124.
- KANWISHER, N. G., & POTTER, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 30-47.
- KARIS, D., FABIANI, M., & DONCHIN, E. (1984). "P300" and memory: Individual differences in the von Restorff effect. *Cognitive Psychology*, **16**, 177-216.
- KEPPEL, G., & UNDERWOOD, B. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior*, **1**, 153-161.
- KLEIN, R. (1993, May). *On the relationships between overt and covert orienting: A view from human performance*. Paper presented at the Third West Coast Attention Conference, Eugene, OR.
- KRAUT, A. G., & SMOTHERGILL, D. W. (1978). A two-factor theory of stimulus-repetition effects. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 191-197.
- KRAUT, A. G., SMOTHERGILL, D. W., & FARKAS, M. S. (1981). Stimulus repetition effects on attention to words and colors. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1303-1311.
- LAMBERT, A., & VOOT, N. (1993). A left visual field bias for semantic encoding of unattended words. *Neuropsychologia*, **31**, 67-73.
- LEVINE, M. (1988). *Effective problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- LEWICKI, P., HILL, T., & SASAKI, I. (1989). Self-perpetuating development of encoding biases. *Journal of Experimental Psychology: General*, **118**, 323-337.
- LOCKSLEY, A., STANGOR, C., HEPBURN, C., GROSOVSKY, E., & HOCHSTRASSER, M. (1984). The ambiguity of recognition memory tests of schema theory. *Cognitive Psychology*, **16**, 421-448.
- LOFTUS, G. R., & MACKWORTH, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 565-572.
- LORCH, E. P., ANDERSON, D. R., & WELLS, A. D. (1984). Effects of irrelevant information on speeded classification tasks: Interference is reduced by habituation. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 850-864.
- LUBOW, R. E. (1973). Latent inhibition. *Psychological Bulletin*, **79**, 398-407.
- MAYLOR, E. A., & HOCKEY, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 777-787.
- MCCLELLAND, J. J., & RUMELHART, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review*, **88**, 375-407.
- MYLES-WORSLEY, M., JOHNSTON, W. A., & SIMONS, M. A. (1988). The influence of expertise on x-ray image processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 553-557.
- NÄÄTÄNEN, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral & Brain Sciences*, **13**, 201-288.
- NÄÄTÄNEN, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- NEELY, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.
- NICKERSON, R. S., & ADAMS, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, **11**, 287-307.
- O'REGAN, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics*, **25**, 501-509.
- POSNER, M. I., & COHEN, Y. (1984). Components of visual orienting. In H. Bouma & D. Bowhuis (Eds.), *Attention and performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- PRIGOGINE, I., & STENGERS, I. (1984). *Order out of chaos*. New York: Bantam.
- REASON, J., & MYCIELSKA, K. (1982). *Absent Minded? The psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ: Prentice-Hall.
- REICHER, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, **81**, 275-280.

- REICHER, G. M., SNYDER, C. R. R., & RICHARDS, J. T. (1976). Familiarity of background characters in visual scanning. *Journal of Experimental Psychology: Human Perception & Performance*, *2*, 522-530.
- SCHMIDT, H. G., & BOSCHUIZEN, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory & Cognition*, *21*, 338-351.
- SCHVANEVELDT, R. W., & MEYER, D. E. (1973). Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 395-409). New York: Academic Press.
- SHAPIRO, K. L., & LOUGHLIN, C. (1993). The locus of inhibition in the priming of static objects: Object token versus location. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 352-363.
- SHARPLESS, S., & JASPER, H. (1956). Habituation of the arousal reaction. *Brain*, *79*, 655-680.
- SHIFFRIN, M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127-190.
- SLAMECKA, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *76*, 504-513.
- SMITH, L. C. (1984). Semantic satiation affects category membership decision time but not lexical priming. *Memory & Cognition*, *12*, 483-488.
- SMITH, L. C., & KLEIN, R. (1990). Evidence for semantic satiation: Repeating a category slows subsequent semantic processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 852-861.
- SOKOLOV, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Psychology*, *25*, 545-580.
- SRULL, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning & Memory*, *7*, 440-463.
- STANGOR, C., & McMILLAN, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social development literatures. *Psychological Bulletin*, *111*, 42-61.
- TIPPER, S. P., BREHAUT, J. C., & DRIVER, J. (1990). Selection of moving and static objects for control of spatially directed action. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 492-504.
- TREISMAN, A. (1992). Perceiving and re-perceiving objects. *American Psychologist*, *47*, 862-875.
- VON HIPPEL, W., JONIDES, J., HILTON, J. L., & SOWMYA, N. (1993). The inhibitory effect of schematic processing on perceptual encoding. *Journal of Personality & Social Psychology*, *64*, 921-935.
- VOSS, J. F., & BISANZ, G. L. (1985). Knowledge and the processing of narrative and expository texts. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 173-198). Hillsdale, NJ: Erlbaum.
- WALLACE, W. P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin*, *63*, 410-424.
- WICKENS, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review*, *77*, 1-15.
- WILSON, F. A. W., O SCALADHE, S. P., & GOLDMAN-RAKIC, P. S. (1993). Dissociation of object and spatial processing domains in primate visual cortex. *Science*, *260*, 1955-1958.
- YANTIS, S. (1993). Stimulus-driven attentional capture and attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 676-681.

NOTES

1. We use the terms *bottom-up* and *top-down* to refer to different directions of flow between lower and upper layers of nodes and *data-driven* and *conceptually driven* to refer to different origins of flow (namely, external input vs. active, upper-layer nodes).

2. The literature on perceptual memory suggests that the bottom-up processing is enhanced for objects that are repeated in the same context (see, e.g., Jacoby, 1983; Jacoby & Dallas, 1981). Thus, it is arguable that the strength of the bottom-up links from the iconic nodes to the upper-layer nodes should be modifiable rather than fixed. We have opted to leave them fixed because we failed to observe perceptual memory effects in our experimental paradigm (Johnston et al., 1993, Experiments 1 and 2). However, modifiable bottom-up connection weights could easily be added to the model and might even broaden its scope.

In addition, either hardwired or modifiable links could be provided between location nodes. We elected to omit such links both because we suspect that it is biologically the most plausible option and because it is computationally the most expedient. However, we tried the other options in earlier simulation runs and observed them producing the basic pattern of novel-popout effects as well.

3. We are currently revising our simulation program so that it can accept the sorts of experimental files needed to conduct studies of word-superiority and perceptual-restoration effects. We will try to simulate these effects, as well as novel-popout effects, using the same set of parameters.

(Manuscript received July 16, 1993;
revision accepted for publication October 22, 1993.)