7

Mechanisms of Inhibition in Long-Term Memory

A New Taxonomy

Michael C. Anderson and Robert A. Bjork

> *The existence of forgetting has never been proved: We only know that some things don't come to mind when we want them.*
>
> Friedrich Nietzsche

I. INTRODUCTION

Most of us have forgotten things that we were sure we knew: The name of a friend or the location in which we safely stowed our airline tickets may, to our embarrassment or horror, simply elude reinstatement. Later exposure to the name or discovery of the errant tickets frequently prompts an *aha!*, because we recognize the name or the event of storing the tickets. Clearly such information often remains in long-term memory, but has been rendered inaccessible. When these retrieval failures are caused by other memory activities, such as the encoding or retrieval of related information, we say that accessibility of the affected items has been inhibited by those activities. More simply, we say the items have undergone *retrieval inhibition*. In this chapter, we consider theoretical mechanisms that might cause retrieval inhibition in episodic memory, with particular emphasis given to the issue of whether inhibitory mechanisms are responsible for impaired recall.

At least two characteristics motivate the choice of the term *inhibition* to describe these memory failures. First, impairment is thought to arise from an active process—such as the learning or strengthening of related

INHIBITORY PROCESSES IN ATTENTION,
MEMORY, AND LANGUAGE

items—that affects specific materials; this impairment contrasts with retrieval failures arising from more global, passive processes, such as changes in the availability of retrieval cues (e.g., as might occur with a change in environmental context) or passive decay. The active, specific nature of these mechanisms can be illustrated with the retroactive interference paradigm employed in classical verbal learning research. In this paradigm, subjects typically learn a first list of verbal paired associates (e.g., Dog–Rock), followed by a second list in which the stimulus terms are the same, but the responses paired with them differ (e.g., Dog–Lamp). On a later cued-recall test, subjects are asked for both responses associated to a given stimulus (e.g., Dog ______ ______). Recall of first-list responses is typically impaired by acquisition of the interpolated list (see, e.g., Postman, 1971; Postman & Underwood, 1973; Crowder, 1976, for excellent reviews). This impairment is dramatically reduced, however, when first-list associates (e.g., Road–Rock) do not share stimulus terms with the second-list associates (e.g., Dog–Lamp). Thus, subsequent memory activity seems to impair recall of earlier responses, and this activity affects some materials much more than others.

Similar impairment can be induced with other procedures, such as part–set cuing. In a common version of this procedure, subjects study a categorized word list (e.g., containing fruits, trees, etc.) and receive an immediate category-cued episodic memory test. Interestingly, if several earlier studied exemplars (e.g., Lemon, Orange) are given along with the category label as cues to help in recalling the remaining words (Cherry, Banana), performance on those remaining words *suffers* rather than benefits, relative to the performance of subjects receiving only the category label (see, e.g., Nickerson, 1984; Roediger & Neely, 1982, for reviews of these and related phenomena). Mere retrieval of information from long-term memory (e.g., Retrieving Orange, given the cue Fruit–Or______) has also been shown to impair related items (but not unrelated items) in both speeded semantic retrieval tasks (Blaxton & Neely, 1983) and in episodic tasks with delays of 20 min (Anderson & Bjork, 1990; Anderson, Bjork, & Bjork, in press). Classic work in repression (Freud, 1952; Erdelyi & Goldberg, 1979), directed forgetting (Bjork, 1972; Epstein, 1972; Roediger & Crowder, 1972), and posthypnotic amnesia (Kihlstrom, 1983) even suggests that subjects can consciously apply such processes, provided that the to-be-forgotten information is in a clearly demarked episode. The important point, for present purposes, is that retrieval inhibition can be induced by activities such as encoding new similar memories or restudying or retrieving existing competitors.

A second characteristic motivating the use of the term inhibition is that subjects' knowledge of the forgotten items can often be clearly demonstrated with recognition tests. In retroactive interference, impairment of first-list responses is dramatically reduced when a multiple choice recognition test is administered (see, e.g., Postman & Stark, 1969; however, see Chandler, 1989, 1993, for reviews of studies in which recognition deficits have been found). In part–set cuing, recall deficits for cued items are small to nonexistent when recognition probability is tested (Slamecka, 1975; Todres & Watkins, 1981), although they can appear with more sensitive recognition procedures (e.g., recognition time—see Neely, Schmidt, & Roediger, 1983; for a related phenomenon, see the fan effect of Anderson, 1974). List strength effects, or the tendency for the strengthening of some items on a study list to impair free recall of the remainder, disappear or even reverse in recognition tests (Ratcliff, Clark, & Shiffrin, 1990). Our own preliminary investigations with retrieval-induced forgetting corroborate all of the preceding findings. Clearly then, much of the impairment in these paradigms is not permanent, because if it were, recognition memory for impaired items would suffer as well. Thus, in retrieval inhibition, learning or strengthening some information inhibits the *accessibility* of related items, but leaves their absolute *availability* unaffected (Tulving & Pearlstone, 1966).

But what exactly does it mean for items to have inhibited accessibility? In its most theoretically neutral sense, this expression simply describes the fact that some activity caused a decrease in recall performance that would not otherwise have occurred. For example, Bjork (1989) emphasized that the term *inhibition* is often used in the memory literature simply "as a descriptor for empirical effects that are the opposite of facilitation" (p. 309). Clearly, these characteristics justify the use of inhibition in this weak sense. However, if what is meant by retrieval inhibition is that an item's representation is suppressed—that is, *the representation's level of activation is reduced by the action of an inhibitory mechanism*—then the term is not clearly justified. As we will illustrate, most empirical phenomena exhibiting retrieval inhibition in the weaker, descriptive sense can be explained adequately without assuming the existence of retrieval inhibition in this stronger, mechanistic sense. The growing interest in inhibitory mechanisms as theoretical constructs (evidenced, for example, by this volume) has made it crucial to have a means of establishing whether inhibitory phenomena truly reflect inhibition in the strong sense.

Assessing whether impaired performance reflects inhibitory processes requires both a familiarity with inhibitory and noninhibitory mechanisms that might produce it, and an empirical criterion for deciding among those mechanisms. This chapter answers these demands by considering such mechanisms and by offering such a criterion. First, we construct a new taxonomy of the various noninhibitory and inhibitory models we have either encountered in the literature or generated on the basis of logical considerations. It is our hope that vivid illustrations of the many theoretical alternatives will be useful as a tool for interpreting empirical findings and will encourage cross-comparison of the reviewed models. Next, we describe a minimal criterion—cue-independent impairment—developed by Anderson and Spellman (1991a, 1991b, 1993) for establishing an effect as inhibitory. Although developed in the context of episodic memory research, the criterion of cue-independent impairment is quite general and may be adapted to most of the domains concerned with the study of inhibition. To illustrate this criterion, we describe a study by Anderson and Spellman demonstrating cue-independent forgetting. This study argues that inhibitory mechanisms produce retrieval inhibition, at least in the case of retrieval-induced forgetting. The final section advances three challenges to those interested in developing theories of the role of inhibition in long-term memory.

For ease of exposition and cross-comparison of models in our taxonomy, we first develop each class of mechanisms in terms of a common experimental paradigm: the retrieval-practice procedure (Anderson & Bjork, 1990; Anderson et al., in press). We discuss other "inhibitory" paradigms as well if results from those paradigms have been offered as evidence for an instance of a given model type. In the next section we describe the retrieval-practice paradigm and the phenomenon that it produces: retrieval-induced forgetting.

II. RETRIEVAL-INDUCED FORGETTING AS A PARADIGM CASE OF RETRIEVAL INHIBITION

In most of the investigations of retrieval inhibition just reviewed, impairment is induced by the learning or strengthening of materials related to the affected items. However, an interesting prediction that follows from current theories of forgetting is that the act of remembering or recalling an item, itself, ought to be a cause of forgetting. It is not that the remembered item should become more susceptible to forgetting; in fact, recalling an item increases the likelihood that it can be recalled again at a later time. Rather, it is *other* items—items that are also associated to the cue or cues guiding retrieval—that should be put in greater jeopardy of being forgotten. Anderson et al. (in press) refer to this possibility as *retrieval-induced forgetting.*

The prediction of retrieval-induced forgetting follows from three assumptions underlying what Anderson et al. call strength-dependent competition models of interference:

(1) The *competition assumption*—Memories associated to a common retrieval cue compete for access to conscious recall when that cue is presented.

(2) The *strength-dependence assumption*—An item's cued-recall performance decreases with increases in the strengths of its competitors' association to the cue.

(3) The *retrieval-based learning assumption*—The act of retrieval is a learning event in the sense that it enhances subsequent recall of the retrieved item.

According to these assumptions, repeated retrievals should strengthen the association of an item to its retrieval cue, rendering other items associated to that cue less accessible. Thus, remembering some things should cause forgetting of others.

To explore the prediction of retrieval-induced forgetting, Anderson et al. devised a new paradigm intended to maximize the negative side effects of retrieval. The procedure, called the *retrieval-practice procedure*, involves three phases: a study phase, a retrieval-practice phase, and a final test phase. In the study phase, subjects study several categories, each composed of several exemplars in category–exemplar format (e.g., Fruit–Orange). After the study phase, subjects engage in directed "retrieval practice" on half of the items from half of the studied categories (three items from each of four categories). Subjects practice retrieving exemplars by completing category-plus-exemplar stem cue tests (e.g., Fruit–Or______) three times for each exemplar, each test interleaved with practice trials on other items. After a retention interval (e.g., 20 min), a final and unexpected category-cued recall test is administered; subjects are cued with each category name and asked to free recall any exemplars of that category that they remember having seen at any point in the experiment.

Of crucial interest in this paradigm is the impact of retrieval practice on the episodic recall of the remaining unpracticed exemplars of the practiced category. According to strength-dependent competition, practiced items should be greatly strengthened, increasing competition for unpracticed exemplars of practiced categories. This increase in competition can be assessed on the delayed cued-recall test for each category in terms of the following logic: If repeated retrieval of competitors impairs remaining unpracticed exemplars, performance on those items should be worse than performance on unpracticed exemplars from baseline categories in which no items were practiced. This outcome is indeed what Anderson et al. found; whereas performance on practiced items improved (as expected) relative to the baseline condition (practiced, M = 74%, as compared to the baseline, M = 49%, an improvement of 25%), such facilitation came only at the cost of performance on unpracticed exemplars of practiced categories (unpracticed exemplars of practiced categories, M = 38%, as compared to the baseline, M = 49%, an impairment of 11%). Thus, incidental strengthening due to retrieval can cause substantial retrieval inhibition. Interestingly, impairment was much greater for (and in some experiments, restricted to) categories composed entirely of high taxonomic frequency (strong) exemplars, than for categories composed of low taxonomic frequency (weak) exemplars.

Two aspects of these findings are worth emphasizing in light of our earlier discussion of retrieval inhibition. First, it is clear from these results that retrieval-induced forgetting is materials-specific in the sense that it is greater for (if not entirely restricted to) unpracticed members of practiced categories than for baseline items which are not in those practiced categories. If retrieval practice inhibited both the critical within-practiced-category items and our within-subjects baseline category items equally, negative effects would be obscured, and we would have no evidence for inhibition. Thus, competition occurs primarily among similar items within a category. Second, these findings appear to illustrate the strength-dependent nature of the competition; as practice enhances performance on practiced items, it impairs performance on the remaining members of the category. Based on these considerations, it seems reasonable to propose an inhibitory process that implements the competitive dynamics among category members.

Although it is natural to suppose that competitive relationships are implemented with inhibitory mechanisms—as they often are in connectionist models—it must be emphasized that the notion of strength-dependent competition is strictly neutral with respect to the implementation of competition. Strength-dependent competition models merely state that there is a quantitative relationship among items sharing a common cue: As some get stronger, *performance* on the others decreases. It is not stated how these performance decrements occur. The mere fact that we obtained the expected retrieval-induced forgetting, and that these findings are consistent with an inhibitory model, does not suffice to demonstrate that such effects are the result of an inhibitory process—especially if a noninhibitory process can be shown to produce the same quantitative relationship. In the next section, we highlight this ambiguity by discussing the many classes of mechanisms, both inhibitory and noninhibitory, by which empirical "inhibitory" effects can be produced. Through this classification, we hope to clarify the differences between inhibitory and noninhibitory models so that they may be distinguished empirically.

III. MODELS OF RETRIEVAL INHIBITION: A NEW TAXONOMY

To decide if a phenomenon truly reflects inhibitory mechanisms, we must have a good characterization of the theoretical alternatives. In this section we consider diverse theoretical approaches to retrieval inhibition, integrating them into a new taxonomy that clarifies the differences between inhibitory and noninhibitory models.

The new taxonomy divides theoretical approaches according to the component of the memory representation that is presumed to have the greatest weight in causing impairment. Although the models that appear in our taxonomy vary in their representational assumptions, all include at least three components: (1) a cue representation, at which memory search begins; (2) one or more target representations, at which memory search may end; and (3) associative links, through which target representations are connected to cue representations. In general, one or more associative links may diverge from a cue representation or converge on a target item. On the basis of these three components, three general sources of impairment may be distinguished, as illustrated in the diagram of our taxonomy in Figure 1: impairment arising from dynamics of the cue representation, the associative links, and the target representation.

In this taxonomy, differing emphases on these three representational components motivate three general classes of approaches, which we will refer to as the *cue bias*, *associative bias* (left side of figure), and *target bias* (right side of figure) models, respectively. This division

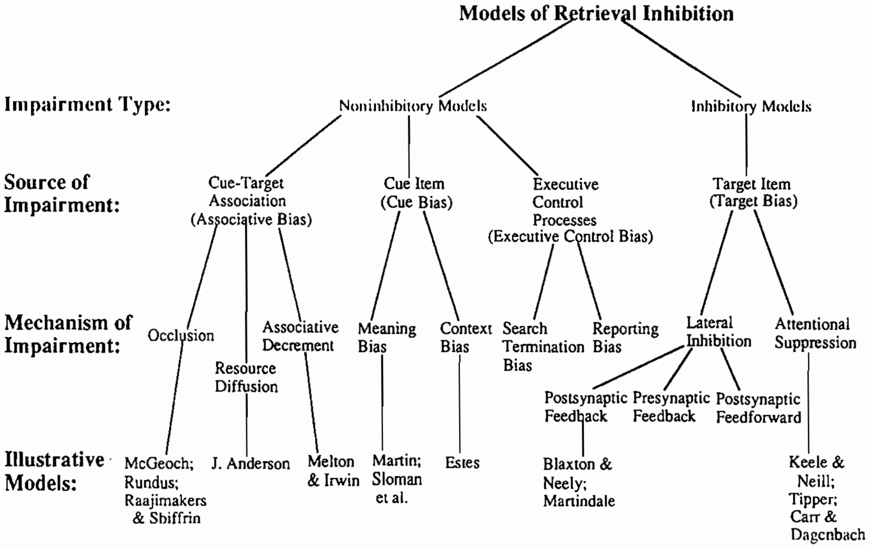

FIGURE 1

A taxonomy of models of retrieval inhibition, organized around the source of impaired performance.

of models into the locus of impairment emphasizes the feature that we argue distinguishes inhibitory models from noninhibitory models: whether the process that causes impairment subtracts activation from the item representation itself. Following this division, Figure 1 depicts target bias models as inhibitory, whereas it depicts cue bias and associative bias models as noninhibitory. An additional noninhibitory source of impairment, *executive control bias*, stems from the manner in which subjects control memory search (middle of figure). Beyond the general locus at which impairment arises, models are further subdivided according to the type of mechanism underlying impairment. For example, associative bias models might attribute impairment to occlusion due to increased associative strengths of competitors, diffusion of finite activational resources across competing items, or to decreased associative strengths between the retrieval cue and the target. Within each of these subclasses, one or more particular instances of that model are depicted.

Although the taxonomy depicted in Figure 1 is concerned with retrieval inhibition in episodic memory, we must emphasize that most of the models considered here are based on general computational strategies that can be applied across many domains. In fact, these models have roots in diverse contexts, including research on classical conditioning (e.g., occlusion and associative decrement models), attention (e.g., resource diffusion, attentional suppression), perception (e.g., lateral inhibition), motor skills (e.g., lateral inhibition), and semantic memory (e.g., lateral inhibition, occlusion). To reflect this generality, we have simplified application of these mechanisms to retrieval inhibition by excluding discussion of particular elements not essential to illustrate a model's basic mechanisms. For example, the representation of temporal context is normally crucial in modeling episodic memory performance, but we exclude it in most cases because it does not contribute to the process being emphasized (with the context bias model as a notable exception).

To emphasize that not all that impairs is inhibitory, we begin discussion of our taxonomy by considering noninhibitory mechanisms. As will be seen, each class of noninhibitory models has phenomena supporting it, suggesting that it is simplistic to seek a unifactor theory of retrieval inhibition that will be able to explain all phenomena. We then complete our taxonomy by discussing inhibitory mechanisms, and how, in some instances, ostensibly noninhibitory models may in fact disguise inhibitory mechanisms.

A. Noninhibitory Accounts of Retrieval Inhibition

Although the term inhibition suggests the operation of inhibitory processes, a surprisingly broad range of noninhibitory mechanisms can accommodate retrieval inhibition. In this section, we describe three general classes of such mechanisms: associative bias, cue bias, and executive control bias models. We review the general logic of each class, distinguish its relevant subtypes, and, when possible, we discuss particular models that illustrate the application of the mechanism to specific inhibitory paradigms.

1. Associative Bias Models One of the most popular styles of explanation for effects of interference and inhibition ascribes the failure to recall an item to decreased efficacy of the retrieval route normally used to access the item. The heavy emphasis of such models on the role of the retrieval route—or the cue–target association—in producing impairment finds its conceptual heritage in traditional verbal learning research and, ultimately, in behaviorist learning theory. In both of these areas the phenomena of crucial interest involve the acquisition and loss of associations between stimuli and responses. Because impaired recall is purported to arise from dynamics in cue–target associations, and not from a subtractive process acting on the item itself, these models are not, by our definition, truly inhibitory.

Many mechanisms could alter the effectiveness of a retrieval route into a given memory item. In this section, we distinguish three types of such mechanisms: occlusion, resource diffusion, and associative decrement. Interestingly, occlusion and resource diffusion models appeal to the positive effects of strengthening related memory items—with no special impairment process—as a means of explaining impaired access to nonstrengthened targets. Associative decrement models, in contrast, assert that strengthening competitors weakens the cue–target associative links for impaired items. For each of these types of model, we illustrate how the effectiveness of the retrieval route into the critical target memory gets reduced, ultimately impairing recall. We then relate each general model to existing theories that, in our view, instantiate them. Where relevant, we describe related empirical phenomena that lend plausibility to the existence of the hypothetical processes.

a. Occlusion The dynamics of occlusion models are vividly illustrated by analogy to the phenomenon of perseveration in monkeys with lesions to the prefrontal cortex. Monkeys with prefrontal lesions often perform acceptably in simple discrimination tasks. When the experimenter reverses the discrimination by reinforcing the previously irrelevant stimulus, however, the lesioned animal persists in selecting the inappropriate stimulus for a duration that varies positively with the amount of training on the original discrimination (Butter, 1969). This perseveration, or the persistent repetition of an inappropriate response, occurs despite lack of reinforcement for that response (Fuster, 1989) and despite previous training on discrimination reversals involving those same stimuli (Butter, 1969). A fairly general phenomenon, perseveration occurs in humans with damage to the prefrontal cortex, and in a variety of other syndromes, in many different forms (Sandson & Albert, 1984). Such phenomena illustrate how competence on a target behavior can be occluded if the organism cannot disengage response production mechanisms from strong alternative response tendencies (e.g., Mishkin, 1964). A similar perseverative process forms the basis for the way in which occlusion models predict effects of retrieval inhibition; failure to recall a target memory occurs because the persistent intrusion of stronger memory targets obstructs production of the target response. Thus, cue–target associations for critical items lose effectiveness not because they are weakened, but because strengthening of competing pathways overfacilitates alternative targets.

To see how occlusion might operate in the context of our retrieval-practice paradigm, consider the simple model of category exemplar retrieval illustrated in Figure 2. In this model, activation spreads from the category retrieval cue to exemplars at a rate that increases with the strength of the category–exemplar association. When an item exceeds a certain threshold level of activation, that node wins access to a limited capacity response production mechanism. In the example illustrated in Figure 2, Orange has been practiced frequently, strengthening the Fruit–Orange associative link (denoted by the thicker black line connecting Fruit and Orange). Because activation spreads more quickly to Orange than to Banana, Orange reaches threshold earlier than Banana and seizes control of the response production mechanism. As long as Orange's greater cue–target associative strength and its heightened activation persist, subsequent attempts to recall Banana will be thwarted by the perseveration of Orange, unless we add a mechanism by which the system can disengage response production (output) from stronger items. Though Banana remains uninhibited (indeed, it may be very active), and though its category–exemplar association remains intact, recall of Banana is impaired.

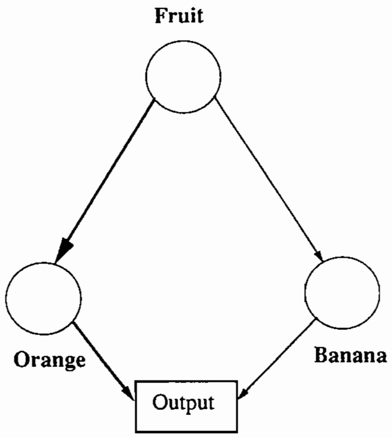

FIGURE 2

Illustration of a simple occlusion model of retrieval-induced forgetting. Strengthening of the Fruit–Orange associative link (designated by the darkened line) leads to capture of a limited capacity output channel.

What reason is there to believe that perseveration underlies retrieval failure in verbal memory, which seems somewhat remote from perseveration in frontal-lobe syndrome? A closer relationship may be found with the perseverative experience that often occurs when we suffer "tip of the tongue" (hereafter, TOT). The analogy between retrieval inhibition and the TOT state is vivid and compelling, and has been espoused both by researchers in the part–set cuing inhibition literature (Reason & Lucas, 1984; Roediger, 1974; Roediger & Neely, 1982) and by advocates of the blocking theory of the TOT state (see, e.g., Brown, 1991, for a review). The subjective experience of TOT includes a feeling that we know the word we seek and, often, that our failure to produce the word stems from the persistent intrusion of a similar, interfering word. Empirical studies of TOT show that subjects rate these persistent alternates as being either more frequently or recently experienced than the target word (Reason & Lucas, 1984). Because the intruding word usually resembles the target, we examine it closely, which seems to make it harder to ignore, further impairing access to the target word. When we engage our attention in some diversionary activity, the correct response often seems to "pop up" (Reason & Lucas, 1984) at a later time. This phenomenon would be consistent with an occlusion mechanism were we to suppose that time and distraction allow us to disengage our attention from the blocking item. If we accept these subjective analyses of TOT, it appears that the characteristics of the TOT experience illustrate the working of an occlusion mechanism in verbal memory.

Though expressed in different terms, variants of this occlusion model have been proposed to account for retroactive inhibition as far back as in the classic work of McGeoch (1936). According to McGeoch's theory of retroactive interference, previously learned responses to a stimulus (e.g., the response Rock to the stimulus Dog) suffer *reproductive inhibition* as a consequence of the acquisition and strengthening of new responses to that stimulus (e.g, the new response Sky to the stimulus Dog). That is, at the time of the testing, response competition from the new, strengthened item (Sky) impairs the reproduction of the earlier learned target response (Rock). Although test performance on the earlier response was said to be inhibited, McGeoch did not believe that the forgotten item was altered or lost. In his terms:

> Inhibition is not thought of as a separate process, but as a function of competition among responses, with a resultant momentary dominance, at least, of one response over another. Responses thus inhibited are not necessarily lost from the subject's repertoire, but are kept by other responses from appearing. (McGeoch, 1942, p. 495)

McGeoch, like modern theorists in TOT research, cited the occurrence of overt intrusions (see McGeoch, 1942, p. 490), as well as the frequency of subjects' reports of covert intrusions as evidence that response competition impaired performance. Thus, the core assumptions of McGeoch's theory of reproductive inhibition—that stronger responses displace weaker responses, though weaker responses remain intact—make it a clear example of an occlusion model.

The basic dynamics of occlusion have also been included in more recent relative strength models of retroactive interference and part–set cuing inhibition. In their extension of the SAM (search of associative memory; Raajimakers & Shiffrin, 1981) model to findings in the interference domain, Mensink and Raajimakers (1988) account for retroactive interference with an occlusion mechanism. In the SAM model, recalling an item requires locating that item in memory (or "sampling" that item, in the language of the model), followed by production of the appropriate response (or "response recovery" of that item). The probability of sampling an item is a function of the strength of association of that item to the current retrieval cue, relative to the strengths of all other memory items associated to that cue (i.e., Strength [Target]/Strength [all associates]). When the cue–target associative strength of an item increases, the relative strengths of all other items associated to that cue decrease, resulting in decreased sampling probability ratios for the nonstrengthened items. According to Mensink and Raajimakers' account of retroactive interference, additional learning trials on the second list strengthen the associations of those new responses to their stimuli, increasing the chances that those new responses will be sampled mistakenly when the older, first-list responses are desired. If subjects mentally perseverate second-list responses when first-list responses are sought, they will ultimately abandon the search, moving on to the next test trial. Essentially the same process model was proposed earlier by Rundus (1973) to account for inhibition from part–set cuing. According to Rundus, presentation of several members of a studied category as retrieval cues strengthened the association of those items to their respective categories. Because the category–exemplar associative strengths of the cue items is presumed to increase, the probability of successfully recalling noncue exemplars should decrease, owing to their weakened relative strength. As in the occlusion model sketched earlier, recall of an item can be impaired simply by increasing the accessibility of competitors.

From the preceding characterization of SAM, it may be difficult to see when and how a limited capacity output channel produces impairment in the model—after all, SAM predicts recall deficits at the stage of *memory sampling.* The stage of *response recovery* is never directly implicated (indeed, the assumption of a limited capacity output channel is rarely made explicit in occlusion models). SAM's status as an occlusion model depends on the interpretation of the sampling process. If sampling is composed of the *activation* of memory targets according to their absolute cue–target associative strengths (as in the model sketched earlier), followed by the *selection* (sampling) of one of those activated targets according to its relative strength of activation, then SAM is a limited capacity output channel (occlusion) model. In this interpretation, the analogue to the capture of the limited capacity output channel by an exemplar (see Figure 2) would be the selection of (sampling) only *one* of the activated targets on which to focus the recovery process. Presumably, this sampling limitation indirectly derives from the restriction that the recovery process can only focus on one response at a time. Thus, if the ratio-rule equation for sampling governs only the *selection* of one among many activated targets, then SAM is an occlusion model.[1]

Although occlusion models are intuitively appealing, especially when illustrated with the subjective experience of TOT states, evidence for such theories warrants caution for at least two reasons. First, there is a tendency to interpret *correlations* between persistent alternates and TOTs as if they were evidence that persistent alternates caused TOTs. It is entirely possible, however, that the reverse is true: A strong correlation between persistent alternates and TOTs is equally good evidence that TOT states cause the generation of such close matches. Given that partial knowledge of the sought-after item often accompanies TOTs (see Brown, 1991, for a review), it is not surprising that the search process often generates close matches, especially when we consider the extreme persistence of people experiencing TOTs. Analogously, subjects in retrieval inhibition paradigms may perseverate in producing strengthened competitors precisely because active inhibition has rendered the desired targets unavailable. Second, even if we grant that occlusion plays a role in causing TOT or retrieval inhibition, we must be cautious not to overextend its explanatory power. For example, it seems likely that blocking cannot be the sole process behind TOT experiences for the simple reason that persistent alternates do not accompany a large

---

[1] As will be seen in our discussion of the resource diffusion model, the ratio rule can be used instead to determine the spread of activation from the cue to the targets, rather than to select targets after they have been activated. Thus, SAM might be interpreted as a resource diffusion model. Because no mention of limited resources is made in discussions of SAM, and because perseveration is stressed, we classified it as an occlusion model.

proportion of the reported incidences in TOT studies. Indeed, whereas the frequency of TOT experiences has been found to increase with age, the incidence of persistent alternates decreases (Burke, MacKay, Worthley, & Wade, 1991). Although occlusion is intuitively appealing, and may be supported by phenomena such as TOT, the empirical case for the role of occlusion processes is not clearly established.

It should be emphasized that in the occlusion model impairment occurs at the stage of final output; limitations on the response channel impede the report of target responses, even though the ability to activate those responses remains intact. If activation of the category cue still activates nonrecalled exemplars, any processing involving the nonrecalled item's representation prior to response production should be unaffected by the strengthening of competitors. For example, nonpracticed exemplars should still be primed by presentation of the category cue, because the amount of activation that spreads to nonpracticed items does not change as a consequence of strengthening competitors (see Yaniv & Meyer, 1987, for work on the TOT state that may support this prediction). This property of occlusion may allow it to be distinguished from the next two classes of noninhibitory mechanisms: resource diffusion and associative decrement.

b. Resource Diffusion The occlusion model asserts that strengthened exemplars impair competitors by seizing control of a limited capacity response–production mechanism. Occlusion requires no special processes that inhibit representations of impaired items or that diminish their cue–target associative strengths. Thus, there should be no decrease in the amount of activation spread to nonstrengthened competitors, despite an increase in the amount received by strengthened exemplars. Indeed, connectionist architectures typically assume that the amount of activation spread by a source node to a particular target is not modulated by the amount spread to other associates. However, other spreading activation mechanisms are possible.

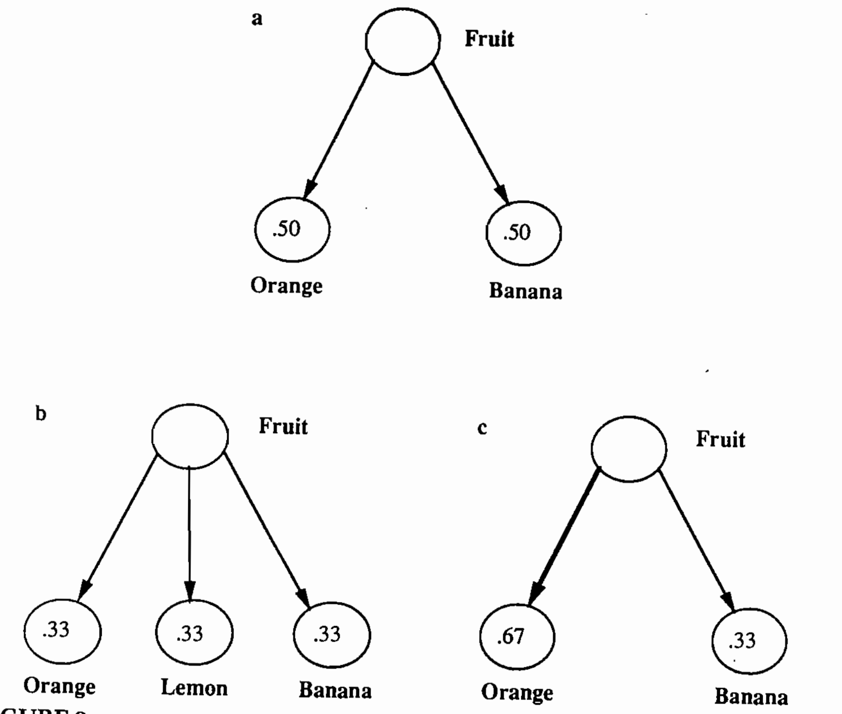

FIGURE 3

Illustration of a resource diffusion model of retrieval-induced forgetting. Strengthening of Fruit–Orange (c) robs Fruit–Banana of activation [note the reduction in percentage activation received by Banana—.33—relative to that received in (a)], as does the addition of a third exemplar (b).

Suppose that activation is a finite resource to be divided among the connections emanating from a node. As the number of elements to which a node connects increases, the amount of activation spread to any given associate should decrease. Furthermore, strengthening the associative links of some of the associates should rob the remaining items of activation because strengthened items should receive more of the total activation. Such decreases in the amount of activation received by nodes should reduce their recall probability, despite the absence of an inhibitory mechanism acting on them. Figure 3 illustrates these effects with our category–exemplar materials. In this example, the rate at which activation spreads along a pathway is determined by the strength of that category–exemplar association, *relative* to the strengths of all associations emanating from the category cue. For this reason, resource diffusion and occlusion mechanisms (Mensink & Raajimakers, 1988; Rundus, 1973) can behave similarly in that the pattern of recall produced by each obeys a general ratio-rule formulation of memory retrieval (Raajimakers & Shiffrin, 1981; see also Luce, 1959, for a more general treatment of "choice" models). Figures 3a through 3c illustrate how resource diffusion produces this behavior. Note that, assuming equally weighted links, the amount of activation received by any given node decreases from .5 of the total in case (a) to .33 of the total in case (b) because activation must be spread among more links. An increase in the associative strength of an existing item results in similar effects, as illustrated in case (c). It must be emphasized that the absolute associative weights linking cues to targets remain unchanged for nonstrengthened items. Decrements in the potential to activate those items arise strictly from the manner in which activation is divided at the source node, enabling the model to capture interference effects without postulating special weight decrement processes. Essentially this mechanism is embodied in the spreading-activation assumptions of Anderson's (1983) ACT* cognitive architecture.

Perhaps the most relevant body of work supporting the notion of activation as a finite resource is Anderson's own work on the fan effect in fact retrieval. The *fan effect* refers to the increase in reaction time that results from an increase in the number of facts associated to a concept. For example, in an early study by Anderson (1974), subjects studied a set of facts until they knew the facts perfectly. Facts took the form of "Person x is in location y" (e.g., subjects might have studied *The lawyer is in the church*). Subjects were then timed to see how quickly they could judge whether they had seen a sentence in the original study phase. Trials with previously studied facts—to which subjects were to respond "Yes"—were interspersed with new facts composed of the same characters and locations, paired differently—to which subjects were to respond "No." Anderson found that subjects took longer to respond "Yes" if they had studied another fact involving the same character or location (e.g., *The lawyer was in the park* or *The doctor was in the church*). Later studies (Lewis & Anderson, 1976) showed that the delay increased with the number of related facts studied (but see also Radvansky & Zacks, 1991, for restrictions on the generality of this finding). Anderson used these findings to argue that the capacity of a concept, such as *lawyer*, to activate associated material is finite and that adding new associates to a concept should diminish its effectiveness in activating any of its associates. Anderson makes this interpretation of the fan effect explicit in his model of that phenomenon (Anderson, 1976).

Various other findings in both short-term and long-term memory, characterized by the cue-overload principle (Watkins, 1978), are broadly consistent with a resource diffusion model. The cue-overload principle states that as a memory retrieval cue becomes associated to more events, the probability of retrieving any one of those events declines—that is, the retrieval cue becomes "overloaded." For example, it is well known that as the number of items on a study list increases, the proportion of items recalled from that list decreases (Murdock, 1960). According to Watkins (1978), recall performance worsens as lists get longer because more items get associated to a general list retrieval cue, overloading it. Consistent with this notion, lists composed of several semantic categories result in dramatically increased recall proportions, presumably because subjects use the less overloaded category names as retrieval cues. Similar characterizations can be given to a variety of the phenomena of interest in this chapter, such as part–set cuing inhibition and retroactive interference, as well as the buildup of proactive interference in short-term memory.

Although a range of data seems consistent with a resource diffusion model of retrieval, those same data can be modeled by most of the other mechanisms described in this chapter. For example, an occlusion model can explain the list length effect if it is assumed that retrieving list items strengthens the associations of those items to the list-context representation. Longer lists, then, by virtue of having more items to be recalled, will have a larger number of previously strengthened (recalled) competitors against which a nonstrengthened remainder must compete. Thus, though it may seem as though the mere presence of additional associates reduces the effectiveness of a list retrieval cue, performance may be impaired because of a higher incidence of retrieval-induced competitor strengthening.

c. Associative Decrement Although it is natural to think that associations both increase and decrease in strength, as they do in connectionist learning schemes, memory theorists frequently assume that associations never go away once learned (see, e.g., the assumptions of Raajimakers & Shiffrin's, 1981, SAM model). This assumption follows from the more general notion that memories are never truly lost, but are merely made inaccessible (e.g., Tulving, 1974; see also, Hintzman, 1978, pp. 297–304; Loftus, 1979; Loftus & Loftus, 1980, for contrasting views). Nonetheless, within the framework of classical associationist learning theory, it is often presumed that decrements in associative strength underlie phenomena such as the extinction of conditioned responses. Recent inquiries into the cellular mechanisms of conditioning render such a decremental mechanism plausible on the neurophysiological level: Synapses in the hippocampus—a brain structure associated with memory storage—demonstrate *associative long-term depression* (the counterpart of long-term potentiation; e.g., Nicoll, Kauer, & Malenka, 1988), a property whereby synaptic efficacy decreases with negative correlations between inputs from that synapse and postsynaptic activity (Stanton & Sejnowski, 1989; see also Levy & Steward, 1983). To the extent that we take such changes in synaptic efficacy to be the neurological underpinnings of associative learning on more complex levels, neural mechanisms exist for both learning through associative increment and *associative decrement.*

Mechanisms of associative decrement can capture the general character of retrieval inhibition. Suppose, as we have for previous models, that learning Banana and Orange results in an associative structure in which the cue (or stimulus) Fruit is linked to the exemplars (or responses). Successful retrieval practice of Fruit–Orange should reinforce that associative link, increasing the probability that Orange will be given again on presentation of Fruit. However, because Banana is also associated to Fruit, there is some probability that it will be retrieved mistakenly (but not reported, because it is inconsistent with the category stem cue Fruit–Or______). The critical assumption for an associative dec rement process would be that such nonreinforced retrieval of Banan would cause a decrease in the associative link between Fruit and Ba nana, akin to the decrease hypothesized to underlie extinction in con ditioning studies. This decrease in associative strength would caus later category-cued recall of Banana to be impaired.

The most widely known example of an associative decrement mech anism—unlearning—was proposed by Melton and Irwin (1940) as pa of their two-factor theory of retroactive interference. Prior to Melton an Irwin, the dominant account of retroactive interference was McGeoch theory of reproductive inhibition (McGeoch, 1936)—a theory that as cribes impaired performance to response competition at test. A natura prediction following from the response competition view is that in creases in overt intrusions of stronger, interpolated material shoul accompany increases in retroactive interference. Melton and Irwin dis covered, however, that as the number of learning trials on an interpo lated list was increased to extreme levels (e.g., 10, 20, or 40 learnin trials), retroactive interference increased, whereas the frequency o intrusions from the interpolated list *decreased.* This dissociatio prompted Melton and Irwin to propose an additional factor contribu ing to interference, a factor "X," tentatively identified as unlearning o the impaired association: "Available alternative hypotheses regardin Factor X identify it with some process which occurs during the lear ing of the interpolated responses and weakens the original S-R rela tionships" (Melton & Irwin, 1940, p. 199). Thus, Melton and Irwin Factor X constitutes an associative decrement model of retroactiv interference.

Associative decrement accounts of retrieval inhibition, such as tha proposed by Melton and Irwin, can be given more precise computationa expression in terms of error-correction learning algorithms currentl popular in connectionist modeling (Rumelhart, Hinton, & William 1986; Sutton & Barto, 1981). The learning rule employed by this class o models, generically referred to as the delta rule (Rumelhart, Hinton, Williams, 1986), specifies the change to an associative weight—eithe positive or negative—on a given learning trial. Figure 4 provides a sim ple version of the delta rule for two-layer associative networks, alon with an illustration of how this mechanism can be applied to th retrieval-practice paradigm. The learning procedure for these model requires presentation of a set of pairs of input and output patterns to th network. Figure 4a specifies the initial input pattern, which, for simpli ity, is an activation input of 1.0 (in a range of 0.0–1.0). This input cause the spread of activation to both Orange and Banana, the weights o which we have (arbitrarily) set to be equal and .5. The spread of activa tion from Fruit generates an output pattern on the units of the outpu

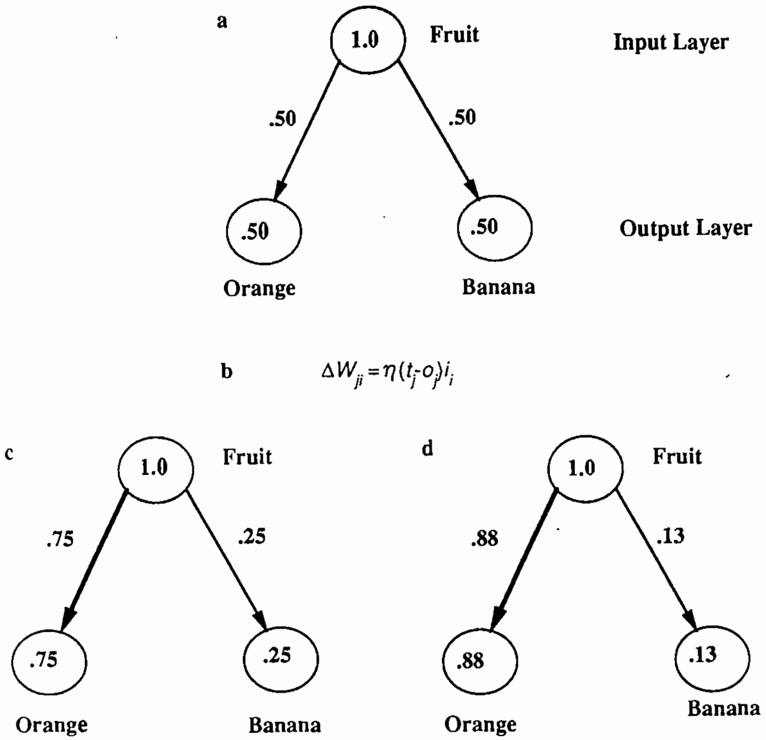

FIGURE 4

Illustration of an associative decrement model of retrieval-induced forgetting. In these examples, the weights into Orange and Banana are written alongside the links, the input (i = 1.0) is written in the category node at the input layer, and the result of spreading activation (the output, o) is written at the output layer exemplars. Weights are set to be equal in (a), and the network is trained for two trials in (c) and (d) to produce (practice) the exemplar Orange by punishing the activation of Banana and rewarding the activation of Orange. This is done according to the equation given in (b), the desired output pattern t = [1,0] and the learning rate, η = .5.

layer (Orange and Banana), a pattern which, in this case, is [.5, .5]. Generation of output is crucial to the learning algorithm, which compares the actual output to the *desired output.* In our example, the model is retrieving Orange, so our desired output pattern has Orange active and Banana inactive: [1, 0].

With these parameters, the delta rule changes weights in a manner that reduces the difference between the actual and desired output: Links decrease for nodes that are too active, whereas links increase for nodes that are not sufficiently active. Changes occur according to the equation in Figure 4b, in which the change of a weight between an item (j) and the category (i), (ΔW_{ji}), is a function of the difference between the target output activation (t_j) and the actual output (o_j), weighted by both a learning rate parameter (η) and the input activation of the sending unit on the input layer (i_j). Figures 4c and 4d illustrate the result of applying this rule to two retrieval practices on the initial representation in 4a, assuming a learning rate parameter, $\eta = .5$. Because Banana is too active (the actual output for Banana in example 4a is .5, not the desired output, 0), its associative weight is decreased, and, because Orange is not active enough, its weight is increased. The same effects occur in 4d, though the magnitude of the changes decreases as actual output gets closer to the target output (changes of .125 as opposed to .250). As in traditional formulations of conditioning, the Fruit–Banana association is punished for causing inappropriate activation of Banana, whereas Fruit–Orange is rewarded.

Although the analogy between extinction and retrieval inhibition is appealing, there are difficulties with interpreting impaired recall performance as reflecting true decreases in stimulus–response (S–R) bonds. If the associative bond between the category and the exemplars were truly damaged, one would expect impaired performance on any task dependent on the integrity of that bond. For instance, the ability to recognize that a response was paired with a given stimulus in a multiple-choice recognition task hinges on the associative bond linking the stimulus to the particular response; according to associative decrement models, recognition memory for items suffering retrieval inhibition should be substantially impaired in this task. Yet, as alluded to in the introduction to this chapter, a central characteristic of retrieval inhibition is impaired recall with relatively intact recognition probability (see, e.g., Slamecka, 1975; McCloskey & Zaragoza, 1985). In those instances in which impaired recognition has been demonstrated (in retroactive interference, see Postman & Stark, 1969; Chandler, 1989, 1993; in part–set cuing inhibition, see Todres & Watkins, 1981), the effects are quite small relative to those observed for cued recall. Enhanced relearning after retrieval inhibition would also seem to cause trouble for the associative decrement approach (although, see Hinton & Plaut, 1987, for an exception)—an observation made even by Melton and Irwin when discussing the merits of the unlearning notion:

> The question remains . . . whether this unlearning factor could be expected to have the functional characteristics assigned to Factor X. . . . [T]he rapid dissipation of the inhibition . . . during relearning requires the assumption that the unlearning of the original responses . . . disappears very rapidly when they are again rewarded. There is no independent evidence on this point. (Melton & Irwin, 1940, p. 201)

Thus, even Melton and Irwin, the progenitors of the associative decrement approach to retroactive interference accepted this hypothesis only tentatively.

In response to such criticisms, one might argue that associations must be decremented more to impair recognition than to impair cued recall. Ignoring the post hoc nature of this rebuttal, more general theoretical analyses of error-correction learning algorithms render such parametric arguments moot. Recent theoretical work by Ratcliff (1990; see also, McCloskey & Cohen, 1989) analyzes the ability of back-propagation (a generalization of the delta rule for multilayer networks) models of learning to capture performance in standard episodic recognition memory tasks. Using a sequential learning procedure (e.g., learn item 1 to criterion, then learn item 2, etc.), Ratcliff (1990) found that error correction models such as back-propagation greatly overpredict the degree of retroactive interference that should be observed in recognition memory tasks. Such catastrophic interference (McCloskey & Cohen, 1989) resulted despite experimentation with a variety of different architectures and learning protocols. In general, these models were best at recognizing the last item trained, and when items trained prior to that were tested, they generated a pattern similar to the final item. Ratcliff (1990) summarizes vividly: "This is like studying the word cat 100 times, studying the word table 100 times, and then finding that cat is not recognized and that when cat is presented for recognition, table is retrieved" (p. 293). Based on the work of McCloskey and Cohen (1989) and Ratcliff's explorations, it appears that such catastrophic interference is a general problem with the ability of error correction algorithms to capture performance in sequential learning tasks. Although there has been some success in reducing catastrophic interference in distributed models of cued recall (see, e.g., Lewandowsky, 1991; Sloman & Rumelhart, 1992), it remains unclear whether these approaches can also model the minimal interference (and occasional facilitation—e.g., Ratcliff et al., 1990; Shiffrin, Ratcliff, & Clark, 1990) observed in recognition memory. To the extent that such algorithms represent the general class of associative decrement models, they fail to provide an adequate account of retrieval inhibition.

2. Cue Bias Models A second major class of noninhibitory models, depicted in Figure 1, focuses on biases in the cues with which subjects search memory. The logic behind these *cue bias* models is simple, and contrasts with that of the associative bias models just reviewed. In both classes of models, retrieval involves activating episodes linked to a specific retrieval cue. According to associative bias models, the efficacy of the link between the cue and the target is diminished; according to cue bias models, the crucial dynamics occur in the *selection of the units or cues* at which activation is initiated: If inappropriate cues are used, retrieval failure will occur. Again, because no special subtractive process acting on target items is proposed, we consider cue bias models to be noninhibitory.

Like occlusion, cue bias might occur in a variety of ways. In this section, we review two types of models in which the cue set is altered to produce retrieval-induced forgetting in the absence of inhibition: meaning bias and context bias models.

a. Meaning Bias In meaning bias models, facilitation of practiced items and impairment of nonpracticed items arise from changes in the representation of the retrieval cue. Specifically, the features constituting the retrieval cue may change either when that cue is used to access certain targets or when it is repeatedly presented with those items. For example, suppose that subjects learned the items Orange, Pineapple, Lemon, and Banana as part of a list of fruits and subsequently performed retrieval practice on the items Orange and Lemon. Such additional practice could bias the representation of the cue, Fruit, toward *Citrus Fruit.* This possibility is illustrated in Figure 5a. Figure 5a depicts a simple semantic network representation of subjects' knowledge of Fruits, including the subcategories Tropical and Citrus. If, during practice of Fruit–Orange and Fruit–Lemon, subjects' default conception of Fruit changed to Citrus Fruit, provision of the nominal cue, Fruit, at test would not allow recall of Tropical Fruit. Using an inappropriate retrieval representation would impair performance on all noncitrus exemplars, not because those items get inhibited, but because they are simply not associated to the functional cue—that is, the cue as it is represented by the subject.

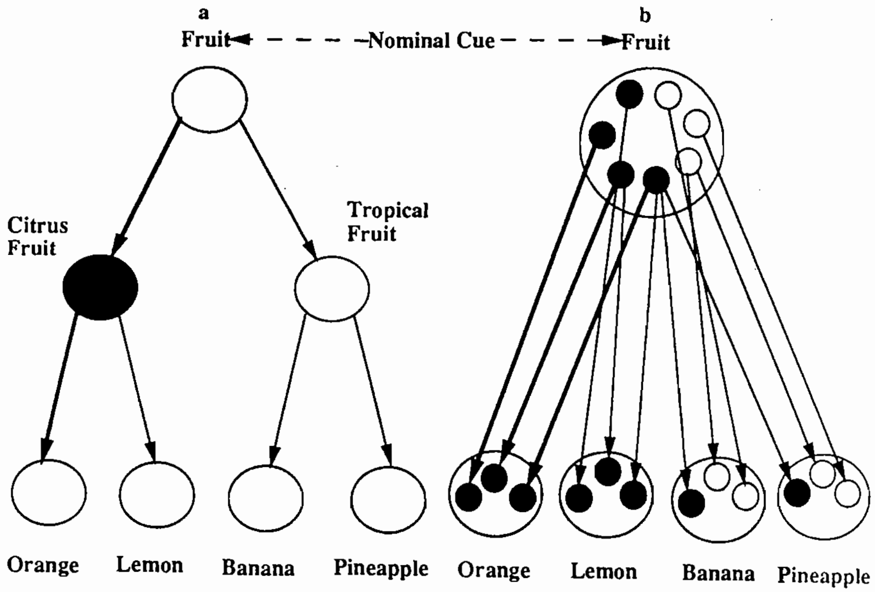

FIGURE 5
Illustrations of meaning bias models of retrieval-induced forgetting. In (a), a localist model, practice of Fruit–Orange shifts subjects' default conception of Fruit to Citrus Fruit (denoted by the darkened node). In (b), a distributed model, practice of Fruit–Orange increases the weights given to citrus features in the retrieval cue (denoted by darkened feature nodes and feature-to-feature-instance links).

Figure 5b depicts a feature-based model that exhibits behavior similar to that exhibited by the model in 5a. In 5b, the concept of Fruit is represented as a constellation of the many features that are associated with Fruit, weighted by their frequency of occurrence in that category. Particular exemplars of Fruit are represented as subsets of instances of those features, associated to their respective features in the general representation. Note that, whereas Tropical and Citrus Fruits do share a common featural element, many more elements are unique to the Tropical and to Citrus subcategories. Retrieval practice would bias the cue representation by increasing the weights of the Citrus features in the general conceptual representation and perhaps by incorporating features unique to practiced items into that representation. An interesting feature of the model in 5b is that it predicts facilitation of nonpracticed items composed of the more heavily weighted features. We will return to this aspect of the meaning bias model at the end of this section.

The meaning bias model asserts that context and the recent processing history of a retrieval cue change the meaning represented for that word on subsequent presentations. This view of the representation of word meaning contrasts with that espoused by traditional theories of natural language semantics, in which the meaning of a word is always represented as a constant set of features (Katz & Postal, 1964; see Barsalou, 1982, for a discussion). An abundance of work, however, illustrates the dynamic, unstable nature of conceptual representations. The overlap in the set of category properties generated by subjects between two sessions separated by a mere week was found to be only slightly more than half (.55—Barsalou, 1987); the speed with which subjects verify whether a property characterizes a category varies significantly as a function of whether an item appears in a neutral or biasing context (Barsalou, 1982) and subjects' typicality ratings for category exemplars vary significantly with linguistic context (Roth & Shoeben, 1983). Such fluctuations in the conceptual representations of words may endure for a long time, as illustrated in an example by Barsalou (1989). Under normal circumstances, seeing the word *frog* may bring to mind the dominant features green and hops. A recent visit to a French restaurant in which frog's legs were consumed, however, may cause one, over the next few days, to encode the property edible upon encountering *frog* because the edibility of frogs has been made salient. Thus, a constant set of features does not enter into the meaning of a word; rather, a word's functional representation varies according to its recent processing history and context. The effectiveness of a retrieval cue, therefore, ought to depend on how previous use changes that cue's meaning.

An early example of how meaning bias could produce impairment is Martin's encoding theory of retroactive interference (Martin, 1971). In Martin's theory, stimulus terms in the typical paired-associate paradigm (e.g., the word *Soap* in the paired associate Soap–Clothes) are assumed to be represented by features varying in saliency to subjects. According to the theory, the features entering into the representation of a stimulus term can change across encodings, depending on how the stimulus term has been used; retroactive interference arises not from response competition, but because the acquisition and strengthening of new responses (e.g., Soap Hand) change which features are salient in the stimulus term, biasing its subsequent encoding at test toward the new response, and away from the original response (in this example, Soap is biased toward bathroom soap and away from detergent). Because, at test, subjects use a different functional encoding of the nominal stimulus term, recall performance on older items is impaired relative to the control group, whose functional encodings of the stimulus terms have not been altered by interpolated learning. Thus, Martin's encoding theory of retroactive interference constitutes a meaning bias model.

Recently, meaning bias has been proposed as a mechanism underlying part–set cuing inhibition (Sloman, Bower, & Roher, 1991). According to Sloman et al., rearranging cues from their original study order changes their interpretation, impairing recall performance on remaining noncue items. Sloman et al.'s view, similar to the earlier strategy-disruption theory of part–set cuing (Basden, Basden, & Galloway, 1977), asserts that this impairment reflects an "incongruency principle"—namely, the more incongruent a "retrieval framework" is with that employed during study, the greater the impairment caused by cuing. Although expressed in different language, Sloman et al.'s (1991) incongruency principle amounts to the dual assertions that (1) recall depends on the overlap between cues at test and the cues during study (encoding specificity) and (2) word meaning depends on context, the latter being evidenced by their application of the incongruency principle to part–set cuing with categorized word lists: "Incongruency could play a role if . . . the interpretation of a category label could be influenced by the presence of category instances that could, for example, make certain attributes of a category more or less salient" (p. 981). Thus, Sloman et al. propose what essentially amounts to a meaning bias model, similar to that outlined here, and by Martin.

Sloman et al. (1991) make the interesting argument that meaning bias can be distinguished from blocking (occlusion) on the basis of the "order-invariance assumption." According to Sloman et al., blocking models such as those proposed by Rundus (1973) and Raajimakers and Shiffrin (1981) predict that the order of presentation of part–set cues should not influence the degree of impairment (order invariance of inhibition). Rearranging cues should not affect the magnitude of impairment because the degree to which cues are strengthened (the crucial determinant to blocking) should not be influenced by their order. In contrast, according to the incongruency principle, cue order should affect interpretation of the cues themselves, enhancing inhibition. To test these predictions, Sloman et al. had subjects listen to a tape-recorded list presenting 36 common names in the order: first name, last name, first name, last name (e.g., Jackie, Smith, Tom, Johnson, Steve, Nelson . . .). Subjects in the congruent cuing condition were presented at test with all of the first names from the first half of the list and all of the last names from the second half (e.g., Jackie, Tom, Steve, . . . Jones, Martin, Robinson) in exactly the same order in which they appeared during study. Subjects in the incongruent order received precisely the same names as did the congruent group, but rearranged so as to form the names of famous persons not on the list (e.g., Jackie, Robinson, Tom, Jones, Steve, Martin . . .). In support of meaning bias, Sloman et al. showed part–set cuing inhibition for recall in the incongruent group, but not for the congruent, although the congruent group showed a trend toward impairment. Order invariance seems to be violated.

Although these results appear to support meaning bias as the mechanism underlying part–set cuing inhibition, the case is unfortunately not clear. First, the property of order invariance applies only to simplistic formulations of blocking processes that ignore the role of interitem associations. The argument for the role of interitem associations in producing blocking is simple. When a cue is presented, subjects retrieve other list items to which that cue was associated during study. If cues are rearranged from the order in which they were originally studied, subjects may acquire new interitem associations among cues that interfere with the use of the previously acquired interitem retrieval routes. Sloman et al.'s materials vividly illustrate how this interitem interference might occur during part–set cuing. Suppose that instead of merely associating items to a general list retrieval cue, subjects in Sloman et al.'s experiment formed interitem associations such that they segmented the study list into first-name–last-name chunks (e.g., the four-item list—Jackie, Smith, Tom, Spellman—would be segmented into two first-name, last-name chunks). When cued with first names from the first half of the list and last names from the second half of the list (e.g., Jackie, Tom, Steve, . . . Jones, Martin, Robinson), subjects probably attempted to retrieve the other name associated to each cue. However, because scrambled cues appeared in first-name last-name format (e.g., Jackie, Robinson, Tom, Jones, Steve, Martin . . .), subjects arguably formed new first-name last-name associations during the cuing process that interfered with recall of previously learned first-name last-name pairs—an occurrence that seems likely because Sloman et al. designed rearranged cues to form names of famous people (e.g., retrieving the item Smith, given the cue Jackie was probably much harder given the cue order "Jackie, Robinson," than the order "Jackie, Tom"). Thus, Sloman et al.'s results are perfectly consistent with a blocking account of inhibition that does not ignore the contribution of interitem associations to recall. The order invariance assumption, therefore, does not distinguish meaning bias from blocking models.

Although order-invariance may not distinguish meaning bias from blocking models, the positive side effects of meaning bias might. As noted earlier, meaning bias models predict enhanced performance for noncued items sharing features with cued items. For example, in the distributed meaning bias model described earlier, practice of Orange and Lemon strengthens Citrus Fruit feature weights, increasing the activation spread to both practiced and nonpracticed items sharing those features. This contrasts with the impairment to nonstrengthened items predicted by both occlusion and resource diffusion models. Sloman et al. (1991) noted the facilitatory prediction of meaning bias models as well: "the extent of part-list inhibition in a category will be negatively correlated with the degree of consensus concerning the set of instances that comprise that category" (p. 981). This prediction must be qualified, however, by the possibility that features of practiced items may also have greater associative strengths to practiced exemplars (as illustrated in Figure 5b by darkened lines). To the extent that feature–exemplar associations get strengthened, feature-level occlusion might offset the benefit to similar, unpracticed items.

b. Context Bias A second subclass of cue bias models, context bias, gives more explanatory weight to fluctuations in retrieval context than to changes in experimenter-provided retrieval cues. Although the meaning of context can be vague, the term generally refers to any feature of a subject's experience, other than the nominal stimulus, that may have occasion to be associated with, and fluctuate independently of the nominal stimulus. For example, thoughts or feelings that a subject may experience while studying a list of words are likely to be incidental to that particular episode, and may well get associated to the study material. Other aspects of context include the subject's environment, sensations (noises, smells), and the subject's understanding of the task. Although difficult to define and control precisely, some notion of context is essential if we are to account for the ability to access items from particular episodes.

A context bias account of retrieval inhibition asserts that recall performance will suffer if the contextual representation used to conduct memory search does not match the one present at encoding. For such a bias to explain part–set cuing inhibition, retroactive interference, and retrieval-induced forgetting, it must further be assumed (1) that the task used to strengthen competing items (e.g., cuing, interpolated learning, retrieval practice) differs enough from the one used at learning to generate a new contextual representation to which only strengthened items become associated; (2) that for the experimental group, the test context matches the context formed during the strengthening process more closely than the learning context; and (3) that the test context for the control group matches the one formed during learning. That is, the test context should be biased toward the strengthening phase, reducing recall of nonstrengthened items relative to the recall for those same items in the nonbiased control condition.

Figure 6 illustrates two variations of context bias applied to our retrieval-induced forgetting task. The model illustrated in Figure 6a is what we call a strategic context bias model; a model in which impairment occurs because the subject actively selects the wrong context in

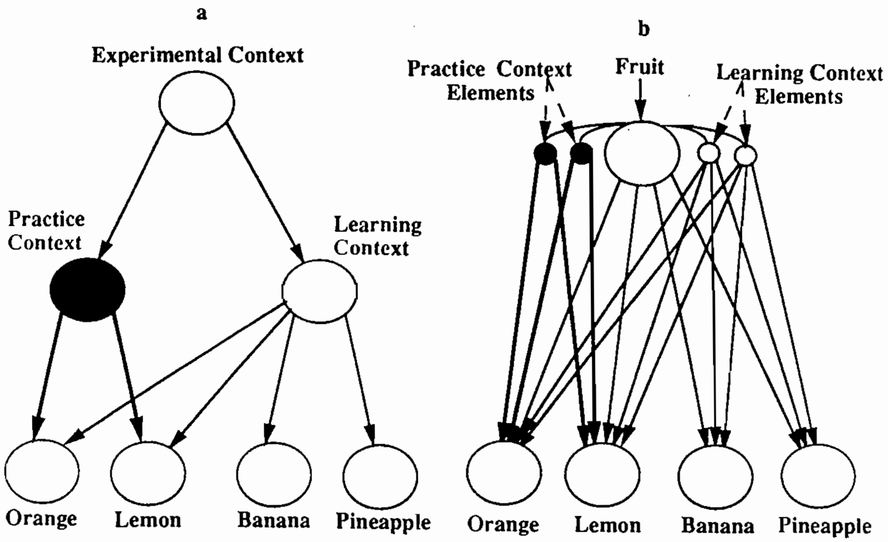

FIGURE 6

Illustrations of context bias models of retrieval-induced forgetting. In (a), a localist model, practicing Fruits biases subjects to use the practice context during search (denoted by darkened context node), causing them to miss items not appearing in that context. In (b), a distributed model, practice context features receive greater weight in the search process (denoted by darkened feature nodes and feature-to-feature-instance links).

which to search. The model assumes that subjects have a representation of the general experimental context, which is further divided into learning and retrieval practice subcontexts. As can be seen in Figure 6a, all experimental items are associated to the learning context, but only those categories and items that were practiced are associated to the practice context. Impaired performance on unpracticed members of practiced categories at final test can then be explained if (1) on seeing a category label, a decision is made about which context to search; (2) selection of context occurs on the basis of familiarity with the category, such that more familiar categories direct search to the practice context and less familiar categories direct search to the learning context; and (3) a switch in context does not occur within a particular category search. On seeing a practiced category, such as Fruit, subjects would search for fruits associated with the practice context, missing items presented only in the learning phase. On seeing an unpracticed category, however, subjects would search with the learning context; because all six items are associated to the learning context, recall should be superior. Thus, according to strategic context bias, impairment results not from inhibition, but from subjects' tendency to look for the right items in the wrong "location."[2]

The strategic context bias model presumes that subjects distinguish among experimental contexts and include these contexts in their conscious search strategies. But if the context to be searched is under the subjects' control, why do they not switch to another context when all items in one have been exhausted? If subjects switched from the practice to the learning context while searching practiced categories, they might retrieve nonpracticed items, eliminating impairment. It is unnecessary, however, to assume that context bias is strategic. Figure 6b illustrates how context bias might impair recall, unbeknownst to the subject. According to automatic context bias, subjects spontaneously associate materials to contextual features with the result that when a category cue is presented at test, the context to which it was most strongly associated is automatically reinstated. Assuming that strengthening phases of inhibitory paradigms associate many additional, unique contextual features to the retrieval cue, presenting a practiced category at test should automatically reinstate the practice context, enhancing recall of practiced items and harming recall of unpracticed items. Unpracticed categories, in contrast, reinstate the learning phase only, enabling access to all exemplars. As in the strategic context bias model, performance is impaired not because items are inhibited, but because the retrieval cues at test are not associated to nonstrengthened items.

---

[2] The authors would like to acknowledge Glen Russ and Andrea Aguiar for the localist context bias argument.

The notion that context is composed of a set of features that automatically fluctuate in saliency was first formalized in Estes's stimulus sampling theory (Estes, 1955). In Estes's theory, the stimulus situation is represented as a population of independently variable components or contextual elements, only a subset of which are active or available at any moment. The set of active elements varies randomly at a rate corresponding to variability in the physical environment and to changes within the organism. To account for performance in learning experiments, the theory asserts that each contextual element is conditioned or associated in an all-or-none fashion to a single response. On a given learning trial, the probability of emitting a particular response is determined by the proportion of the sampled contextual elements associated to the response. When a response is emitted and reinforced, the entire set of sampled elements becomes associated to that response. Other responses previously conditioned to those elements lose their associations.

Stimulus sampling theory accounts for retroactive interference in the following manner. During acquisition of the initial list of paired associates, the currently active contextual elements (including the sampled elements of the stimulus and of the context) become associated to responses. Acquisition of the interpolated list is then accompanied by three types of changes: first, newly active contextual elements get associated to items on the new list; second, some elements associated to first-list items fluctuate out of the available set; and third, contextual elements associated to first-list responses that remain active during interpolated learning are unlearned. If memory tests for first-list responses immediately follow interpolated learning, performance is impaired because most of the active contextual elements are associated to the interpolated list. As time passes, however, and more contextual elements fluctuate in and out of the active set, first-list responses spontaneously recover because some associated contextual elements are reinstated. Thus, as in the model outlined earlier, impairment occurs because the test context is more related to that active in the strengthening phase. Recent accounts of interference phenomena (Mensink & Raajimakers, 1988) have used similar fluctuation mechanisms, although these theories do not stress contextual fluctuation as the primary mechanism underlying retroactive interference.

Although automatic context bias models stress different aspects of the stimulus situation in their explanation of inhibitory phenomena, they may rely on the more basic blocking or finite resource assumptions discussed earlier. This dependence can be illustrated by examining the two ways in which automatic context bias might impair nonstrengthened items. First, it can be assumed—as it often is in stimulus sampling theory—that the number of contextual features in the active set must be a constant size. The strengthening phase of an inhibitory paradigm might then impair nonstrengthened items by adding many new unique features to the active set, forcing out the contextual elements most useful in retrieving nonstrengthened items. Such an approach constitutes a finite resource model of retrieval inhibition, where the size of the active set is the limited resource and where the greater number of contextual elements for strengthened items is like a strengthened association. Second, the number of elements that may be part of a retrieval cue may vary. Practicing some items would then add new, unique elements to the active set, increasing recall probability for strengthened items. For the recall probability of nonstrengthened items to decrease as a result of an increase for strengthened items, some form of occlusion mechanism must be assumed. Thus, context bias may not be a basic mechanism underlying retrieval inhibition in the sense that it may be derivable from other mechanisms.

3. Executive Control Bias The third major class of noninhibitory models illustrated in Figure 1—executive control bias—emphasizes biases in the manner in which subjects execute the search process. So far, all explanations of retrieval inhibition have attributed decreases in recall to impaired item access. However, performance impairments can also result from decisions the subject makes, such as when to start or stop searching, what cues to emphasize and for how long, and whether or not to report a retrieved item overtly. In this section we sketch two ways that impaired recall might result from such executive control biases: search termination bias and reporting bias.

a. Search Termination Bias Subjects' final performance on a recall task depends on when they choose to terminate memory search. The choice of a stopping criterion might lead to impairment of nonstrengthened items in several ways. The first, and perhaps most obvious way would be if subjects underestimate the number of items in the experimental condition, or if they overestimate the number in the control condition. If subjects underestimate the number of items in the experimental condition, search may terminate before all accessible targets have been retrieved. Conversely, if subjects overestimate the number of items in the control condition, search may proceed longer than normal, enabling subjects to retrieve more difficult targets. In either case, recall differences will arise that are not the result of inhibitory mechanisms.

Set-size underestimation seems especially likely in assessments of part–set cuing inhibition in semantic memory. Unless the target set has a well-known size (e.g., the 50 states), subjects' estimates of it may vary considerably, depending on their conception of how the set is defined. For example, the target set size estimate for the category Soaps might be smaller given the examples Ivory, Dial, and Dove, relative to the estimate given the examples, Ivory, Tide, and Joy, because the latter example set broadens the target set definition to include laundry and dish soaps. To the extent that provision of part-set cues narrows subjects' conception of the target set relative to that ordinarily generated without cues, performance will be impaired in the cued condition. Although this example of set-size bias can be regarded as a form of meaning bias, the bias in the meaning of the set has its impact on the point at which subjects voluntarily terminate search rather than on subjects' ability to retrieve exemplars at all.

Set-size estimates might also cause impairment if subjects underestimate the target set size in both the baseline and the experimental conditions. For example, suppose subjects studied 20 exemplars from a category, but one subject estimated the total number of items as having been 15. Consider what might happen in an experimental condition in which 10 of the 20 exemplars are strengthened in some way. If the subject estimates there to be 15 items in total and recalls all of the strengthened subset, search will terminate after retrieving only 5 items of the nonstrengthened subset, although some of the remaining 5 nonstrengthened items may be accessible. When no additional strengthening occurs, subjects may be equally likely to retrieve exemplars from each of the two subsets of the category. If subjects again only retrieve the estimated 15 items, an average of 7.5 items should be recalled from the noncued subset, yielding an advantage of 2.5 items over the experimental condition. Thus, the manner in which subjects control search processes can yield patterns that appear to reflect inhibition, though they do not.

b. Reporting Bias A second executive control factor that might contribute to an apparent inhibitory effect is the nature of the reporting criterion adopted by the subject. It is clear that recalling items in an episodic memory experiment requires the discrimination of items actually presented in the target episode from similar items that did not appear. For example, an episodic memory experiment using categorized word lists requires the exclusion of exemplars not studied. It seems reasonable to expect that subjects often recall more than they report, and that a certain proportion of items recalled, but not reported, are items that were presented in the experiment that have been inappropriately classified as nonpresented items (see, e.g., Tulving & Thomson, 1973, for a demonstration of the phenomenon of recognition failure of recallable words). If such discrimination failures occur more frequently in the experimental condition than in the control, the result might appear to be inhibition.

There are many ways that reporting bias might arise. For example, suppose that the strengthening phase of an inhibition experiment induced a bias in the contextual representation similar to that outlined in our discussion of context bias models, but that context bias was not sufficiently potent to cause retrieval access failures. Nonetheless, subjects may rely on the match between contextual information stored with a memory item and that which is present in the retrieval cue to assess whether an item was presented. If the category cue is biased toward the strengthening context, nonstrengthened items may not be reported because mismatches in contextual information may lead subjects to be underconfident in their assessment of list membership. Such difficulties would not arise for baseline categories, for which the match between the contextual information present in the retrieval cue and that present in the item would be much stronger. Thus, an apparent inhibitory effect might arise merely from biases in subjects' inclination to report retrieved items.

Reporting bias might also produce false inhibitory effects if the recall test directs (either explicitly or implicitly) subjects to retrieve a smaller number of responses than are actually available. Limitation on the number of responses forces subjects to choose which are the most appropriate to report, and depending on subjects' criteria, false inhibitory effects may result. For example, this factor may contribute to the much debated misleading information effect (Loftus, 1979) in eyewitness memory experiments. In the misleading information paradigm, subjects choose which of two photographic slides they recognize as having been presented earlier. One is the originally viewed slide and the other is a similar slide with a detail altered (e.g., if the original slide depicted a car running through a stop sign, the altered slide may depict the same car running through a yield sign). The misleading information effect refers to subjects' tendency to choose the altered slide if, after viewing the slides, they received a verbal summary containing misleading information (e.g., including the inaccurate detail that the car ran through the yield sign). Loftus interprets the misinformation effect as evidence for alteration of the original memory trace, but critics have argued that the effect can arise even when subjects remember both the original slide and the misleading information (Lindsay, 1993; McCloskey & Zaragoza, 1985). This might occur because (1) the task implies that there is only one right answer, and (2) subjects choose the wrong answer because they assume that the experimenters would not make a mistake about their own materials. Thus, apparent memory deficits can arise purely on the basis of what subjects choose to report. Similar arguments about response competition at test motivated the switch from modified free recall tests (MFR tests—recall tests which required that only one response be given to the stimulus term of a paired associate) to modified-modified free recall tests (MMFR—recall tests which required that all previously learned responses be given to the stimulus term of a paired associate) in classical verbal learning research (Barnes & Underwood, 1959).

B. Inhibitory Models of Retrieval Inhibition

As the previous section demonstrates, most long-term memory phenomena described as inhibition can be explained without recourse to theoretical inhibitory processes. These models are attractive in part because of their theoretical parsimony, capturing "inhibitory" behavior without additional mechanisms. The existence of noninhibitory alternatives, however, does not rule out the possibility that inhibitory mechanisms are the primary cause of retrieval inhibition, nor does it nullify the value of considering how true inhibitory mechanisms might generate such effects. In this section, we consider how inhibitory mechanisms might produce retrieval inhibition.

1. Target Bias Models As emphasized in the introduction to this chapter, we reserve the theoretical term *inhibition* for performance impairments produced by decreases in the level of activation of the target item arising from the action of an activation-reducing (inhibitory) mechanism. Because the target item representation is affected, inhibitory models can be classified as *target bias* models according to the taxonomy depicted in Figure 1. Although it is common for researchers to appeal to such target inhibition in discussions of retrieval inhibition, relatively few explicit inhibitory models have been proposed. In this section, we explore and elaborate on two broad classes of inhibitory models of retrieval inhibition: lateral inhibition and attentional suppression. Although each class of model captures the core phenomena using decrements in activation to the target, the approaches are sufficiently distinct so as to lead to different ways of thinking about retrieval inhibition.

a. Lateral Inhibition The mechanisms of lateral inhibition, as they are employed in computational models of cognitive processes, are based on an analogy to the ubiquitous mechanism of lateral inhibition in the nervous system. *Lateral inhibition* refers to phenomena whereby one neuron, A, inhibits another neuron, B, usually via a third (inhibitory) interneuron connecting the two. The circuitry is referred to as lateral inhibition because the neurons that inhibit one another usually are elements of pathways that run in parallel. In computational modeling, lateral inhibition is used to ensure that two processing elements do not remain active simultaneously, if, for reasons of physical or logical incompatibility, their concurrent action is undesirable (e.g., if two elements require a common effector system, paralysis might result if one of the elements were not disabled). Lateral inhibition has been used to model a variety of phenomena, including interference in selective attention (Walley & Weiden, 1973), lexical disambiguation in language comprehension (McClelland & Rumelhart, 1981), phrasal disambiguation in parsing (Waltz & Pollack, 1985), the recall of sequences of items from short-term memory (Estes, 1972), and the production of letter sequences in typing (Rumelhart & Norman, 1982).

Before we consider specific applications of the concept of lateral inhibition to retrieval inhibition, it is useful, as a source of theoretical analogies, to consider the different ways in which lateral inhibition is manifested physiologically. On a physiological level, lateral inhibitory circuits can be classified according to several orthogonal dimensions. First, inhibitory circuits can either be feedback (Figures 7a and 7c) or feedforward (Figures 7b and 7d) in form. Figures 7a and 7c illustrate a feedback circuit in which inhibition occurs between elements in the parallel pathways. The term *feedback inhibition* is used because the inhibitory interneuron acts on a prior point (or at the same point) in the neuronal pathway. Figures 7b and 7d illustrate *feedforward inhibition*, in which inhibition acts on elements further ahead in the pathway. The second dimension along which circuits may be classified concerns whether or not inhibition occurs presynaptically or postsynaptically. In postsynaptic inhibition (Figures 7c and 7d), the inhibitory interneuron synapses on the postsynaptic cell itself, hyperpolarizing or inhibiting it, making it less responsive to any excitatory input. Such postsynaptic inhibition contrasts with presynaptic inhibition (Figures 7a and 7b) in which the interneuron synapses on the axon of the presynaptic cell, rather than on the postsynaptic neuron. This axo-axonic (axon-to-axon) synapse reduces the ability of the excitatory synapse to depolarize the postsynaptic cell, effectively gating the input to that cell from that particular connection. It is important to emphasize that the postsynaptic cell remains uninhibited, which, unlike in postsynaptic inhibition, leaves that neuron largely unaffected with respect to excitation from other synaptic connections. So, in considering lateral inhibitory architectures for retrieval inhibition, reference can be made to at least four basic circuits.

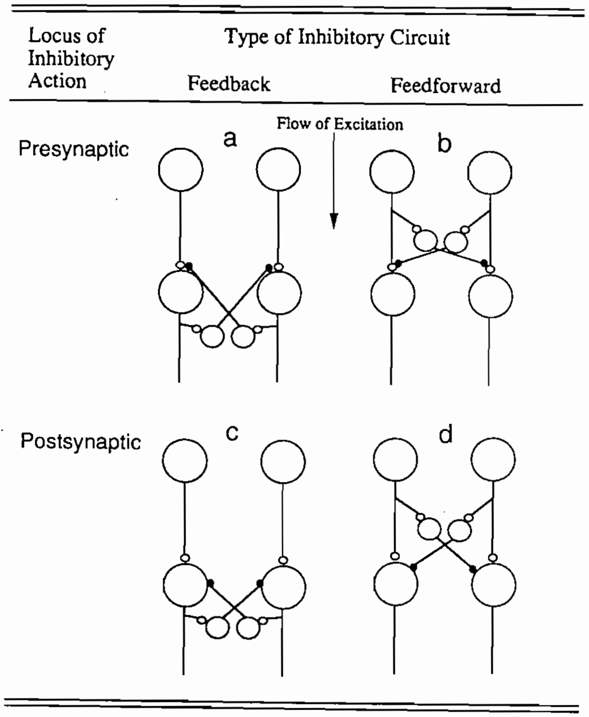

FIGURE 7

Lateral inhibitory circuits, classified according to whether they operate presynaptically (a and b) or postsynaptically (c and d), and to whether they are feedforward (b and d) or feedback (a and c) in design. Solid circles indicate inhibitory synapses and open circles indicate excitatory synapses.

In this section, we sketch three simple models of retrieval inhibition based on analogies to the circuitry just reviewed. In all cases, we use our retrieval practice paradigm to illustrate the operation of the lateral inhibitory mechanism, but the functional properties of these as memory models are not restricted to this paradigm. The first model applies the postsynaptic feedback circuit and entails linking category exemplars with lateral inhibitory connections. It is essentially this model that is mentioned by the few authors who have explicitly proposed (or at least, alluded to) inhibitory models of retrieval inhibition (Blaxton & Neely, 1983; Martindale, 1981; Roediger & Neely, 1982). The remaining two models are based on presynaptic feedback and postsynaptic feedforward circuits. To our knowledge, no one has proposed anything like these models in the domain of retrieval inhibition. They are sketched mainly to illustrate how inhibitory models might capture the behavior of ostensibly noninhibitory mechanisms such as occlusion and resource diffusion.

i. Postsynaptic Feedback Model of Retrieval Inhibition In the preceding postsynaptic feedback circuit, neurons at the same level in parallel pathways inhibit one another via inhibitory interneurons. A similar lateral inhibitory circuit can model the negative interdependencies among category exemplars in retrieval inhibition. Figure 8 illustrates this circuit with our retrieval inhibition stimuli. As in the noninhibitory models reviewed previously, categories and exemplars are represented hierarchically. Retrieval occurs by a spread of activation

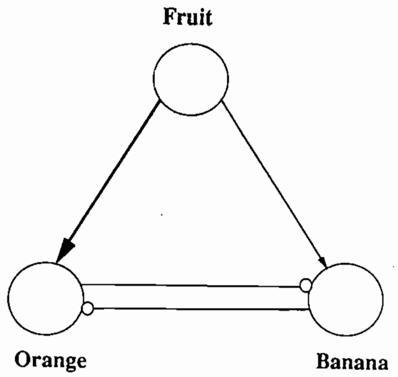

FIGURE 8

A postsynaptic feedback model of retrieval-induced forgetting. Links ending with circles represent inhibitory connections between units; links ending with arrows represent excitatory connections.

from the category cue to exemplars, ending with the production of a single response from the set of activated alternatives. Like the occlusion model reviewed earlier, only one item can be reported per retrieval attempt because of a limited capacity output channel. Unlike the occlusion model, however, simultaneous activation of many alternatives causes difficulty in response resolution. Lateral inhibition enters the scene to resolve this response competition, providing feedback to enhance differences in activation across exemplars. For instance, given a slight activation advantage of the exemplar Orange over Banana, more inhibition will flow from Orange to Banana than from Banana to Orange. As Banana gets inhibited, its ability to inhibit Orange decreases further, increasing the competitive advantage enjoyed by Orange. Thus, a small amount of evidence in favor of one item is self-reinforcing. After one alternative reaches threshold, a response is produced and the recall process begins again.

Inhibition of the item Banana caused by additional retrieval practice of Orange can be modeled in a number of ways. The simplest approach would assert that Orange, because of its strengthened category–exemplar association, gains a consistent early activation advantage over Banana. As detailed previously, negative feedback via lateral inhibitory links would magnify this advantage, causing the strongest items, such as Orange, to be recalled early. Remaining items would have an order and probability of recall corresponding to their strength of association to the category cue. A more complex approach would assert that retrieval practice has a lasting impact on unpracticed items. For example, repeated inhibition via lateral inhibition might decrease the resting level of activation for unpracticed exemplars, or alternatively, increase response threshold. In either case, the inhibited state of unpracticed exemplars would diminish the effect of activation spread from the category cue, decreasing recall probability for as long as such changes were in force. We discuss the issue of the durability of inhibitory effects further in the final section of this chapter.

ii. Variants of the Lateral Inhibitory Model Although lateral inhibition among category exemplars is the most straightforward inhibitory model of retrieval inhibition, other architectures can achieve similar ends. In this section, we sketch two models based on two of the alternative circuitries reviewed previously: presynaptic feedback and postsynaptic feedforward circuits. These alternative models illustrate how many ostensibly noninhibitory models may be recast in terms of lateral inhibition.[3]

The observant reader may have noted during our discussion of presynaptic inhibition that presynaptic inhibition has properties similar to those presumed in a resource diffusion model of retrieval inhibition. In presynaptic inhibition, the ability of a pathway or axon to activate its postsynaptic neuron is disabled, though the pathway itself is not damaged. The postsynaptic neuron is not hyperpolarized, leaving that neuron free to be activated by alternative excitatory connections. Similarly, impaired recall in the resource diffusion model results from decreased activation flow into the impaired item from a particular link. The link itself is not damaged or unlearned, nor does the exemplar node suffer a subtractive effect. Thus, the behavior of a resource diffusion model can be captured by a presynaptic feedback circuit, such as that illustrated in Figure 9.

In Figure 9, the elementary representation used in many of the models reviewed thus far has been augmented to form the computational analogue of a presynaptic feedback circuit. First, the simple connections between the cue and the target have been replaced by link nodes that, like other nodes in the network, can be activated or inhibited. These link nodes play a role similar to that of the synaptic button in presynaptic inhibition. Second, links emanating from the same retrieval cue compete with one another via lateral inhibitory links. These lateral inhibitory links, like axo-axonic synapses in presynaptic inhibition, reduce the transmission ability of competing associative pathways. During retrieval, activation spreads from the retrieval cue through link nodes to the targets. Meanwhile, link nodes inhibit one another,

---

[3] The models discussed in this section do not fit the current definition of target bias because structures other than the item representation (e.g., link nodes and response nodes) are inhibited. These models are discussed in this section primarily for convenience. However, a more general notion, such as "node bias" might categorize these accounts together with the current target bias models.

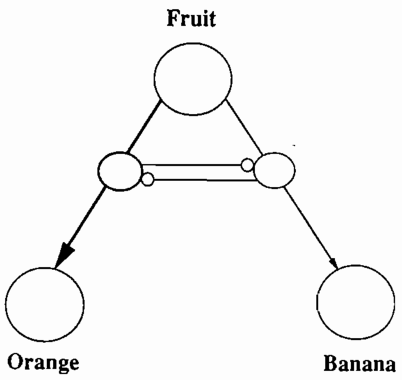

FIGURE 9

A presynaptic feedback model of retrieval-induced forgetting. Links ending with circles represent inhibitory connections; links ending with arrows, excitatory connections. Smaller circles on links between the cue and the target represent "link" or relational nodes that compete with each other through lateral inhibition. Inhibition of the Fruit–Banana link node impairs the spread of activation to Banana.

reducing their effectiveness at propagating activation. This gating mechanism captures cue-overload effects as well as the negative effects of strengthening competitors: as the number of items associated to a retrieval cue goes up, the inhibitory input transmitted to a link node from other link nodes increases, reducing the ability to activate targets. Furthermore, as practice strengthens a link, that link node's ability to inhibit competitors increases, reducing the ability of those competitors to activate their targets. As with the resource diffusion model, however, the targets themselves remain accessible from other retrieval routes. Thus, the circuitry of lateral inhibition can be used in a gating fashion to accomplish the functions of a resource diffusion model.

Like the resource diffusion model, lateral inhibition can capture the behavior of the occlusion model. In occlusion, nonstrengthened items are not themselves impaired, nor are they less capable of receiving activation from the category retrieval cue. Instead, performance on nonstrengthened items is impaired because stronger competitors seize control of a limited capacity output channel. Although such a model appears to be noninhibitory, this claim may be criticized in that it glosses over exactly how a response channel comes to have limited capacity. That is, it can be argued that such an assumption merely disguises or renames an inhibitory mechanism at the level of response production. To illustrate this point, Figure 10 shows how a postsynaptic feedforward network might implement a limited-capacity response channel. In the model illustrated in Figure 10, activation of a category exemplar excites the response associated to that exemplar and inhibits those responses associated to competing memory targets. This feedforward inhibition from exemplars to competing responses implements a limited-capacity output channel. This circuitry has the effect that items more strongly associated to a retrieval cue will be more effective at inhibiting responses for competing items. Like the occlusion model, the memory target itself may remain active, despite impairment of its associated response. Thus, a feedforward lateral inhibitory model can allow strong responses to exhibit behavior resembling the capture of a limited-capacity output channel.

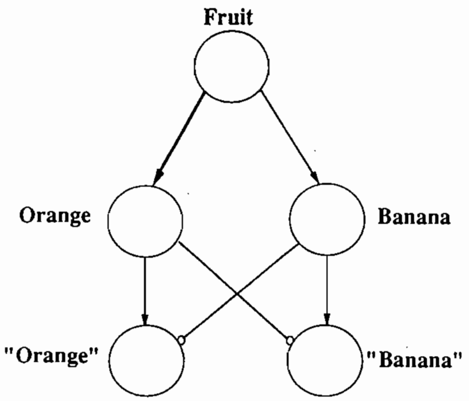

FIGURE 10

A postsynaptic feedforward model of retrieval-induced forgetting. Lexical units representing the verbal response for each exemplar have been added to the model. Feedforward inhibitory links to these lexical units implement a limited capacity output channel.

b. Attentional Suppression A second class of memory inhibition models draws on an analogy between selective attention and memory retrieval. Selective attention and memory retrieval tasks are similar in that both require the production of a response, the appropriateness of which depends on the isolation of a mental representation of one object or piece of information from among a set of activated alternatives. For example, in the Stroop task—a classic paradigm for investigating selective attention—subjects must produce the name of the color in which a word is printed, ignoring the word itself. If the printed word denotes a color, subjects must isolate which of the two activated color names refers to the ink color in order to give the correct response. In memory retrieval, the cues provided—because they tend to be underspecified—may activate many candidate items before the final target is selected for report. Retrieving the appropriate item requires the discrimination of that item from inappropriate competitors, as in the Stroop task. Thus, isolating a target memory from among a set of activated alternatives requires the solution of computational problems similar to those encountered in isolating the memory representations of task-relevant objects in selective attention. In the selective attention domain, it has been proposed that attentional processes assist response selection both by enhancing a target's activation and by decreasing the activation of alternatives through inhibition (Keele & Neill, 1978; Neely, 1977; Neill, 1989; Neill & Westberry, 1987; Neumann & DeSchepper, 1992; Tipper, 1985; see also Gernsbacher, Barner, & Faust, 1990, for a similar mechanism applied in the context of language comprehension). If retrieval can be regarded as the internal focus of attention on memory items and if inhibition deactivates nontargets in selective attention tasks, inhibition may deactivate memory competitors, resulting in retrieval inhibition.

To see how attentional suppression could produce retrieval inhibition, consider how suppression might occur in our retrieval-practice procedure. During retrieval practice, subjects are cued, one at a time, with the name of a category and the first two letters of one of its exemplars. It can be argued that presentation of a category cue activates associated exemplars, particularly those studied recently. To produce the correct response, however, the activated set must be narrowed to the one item that fits the stem cue. The crucial assumption is that focusing search onto the target memory occurs by mechanisms similar to those proposed for attentional focusing: active enhancement of the target exemplar coupled with active inhibition directed at interfering nontarget exemplars. Because, over successive practices, unpracticed exemplars of practiced categories appear repeatedly in the set of rejected alternatives, they are inhibited on later tests of recall. Exactly such a mechanism was proposed by Blaxton and Neely (1983) as one account of their semantic inhibition findings, and by Anderson, Bjork, and Bjork (in press).

The role of inhibitory attentional mechanisms in retrieval was recently emphasized in work by Dagenbach, Carr, and Barnhardt (1990) and Carr and Dagenbach (1990) on retrieval from semantic memory. Dagenbach et al. exposed subjects to the definitions of extremely low-frequency English words for 15 sec each. Of crucial interest was subjects' performance in a priming task that followed the vocabulary acquisition phase. The newly learned vocabulary words served as primes for well-known target words that were either semantically related or unrelated. Subjects were instructed to try to recall the prime's meaning when it appeared, and then to respond to the target as quickly as possible. When subjects succeeded in recalling the prime's definition, as determined by performance on a recall task administered after the lexical decision phase, lexical decisions on related target words were faster relative to performance on trials in which the prime and target were semantically unrelated; surprisingly, when subjects failed to recall the definition, they were slower to make a lexical decision on the related target. Dagenbach et al. attributed impairment of semantically related targets to the action of an attentionally based center-surround retrieval mechanism, analogous to the center-surround mechanisms proposed in theories of attention (e.g., Walley & Weiden, 1973). As suggested in Dagenbach et al., and detailed by Carr and Dagenbach (1990), this mechanism assists retrieval of weakly activated items by enhancing desired targets (the center) while inhibiting more strongly activated items in the surrounding semantic space (the surround). Thus, Dagenbach and Carr's center-surround mechanism, like the attentional suppression process sketched previously, attributes impairment to an active attentional process that enhances discriminability of items in memory.

One might ask whether attentional suppression truly differs from lateral inhibition, because both emphasize the deactivation of interfering competitors. Lateral inhibitory theories have the further advantage, over attentional suppression, of expression as an explicit computational mechanism. There are virtues, however, to viewing retrieval inhibition in terms of attentional suppression. The metaphor of retrieval as the internal focus of attention brings to retrieval processes all of the characteristics that attention has in its original theoretical context. First, it has an active, goal-directed character, diverging from the automaticity implied by traditional spreading activation theories of retrieval (see, Bjork, 1989, for a related argument). Lateral inhibitory mechanisms are also often taken as a passive, noneffortful, automatic form of competition among similar units (e.g., Norman & Shallice, 1980). However, if retrieval inhibition results from an active attentional suppression process, we might expect less impairment for special populations whose deficits supposedly arise from a decreased ability to apply attentional inhibition (e.g., frontal lobe damaged patients: Fuster, 1989; Luria, 1966, pp. 218–295; Mishkin, 1964; Shallice, 1988; schizophrenics: Beech, Powell, McWilliams, & Claridge, 1989; senior citizens: Hartman & Hasher, 1991; Hasher and Zacks, 1988; Hasher, Stolzfus, Zacks, & Rypma, 1991; see also Dempster, 1991, for a discussion of the role of controlled inhibition in general intelligence).

Second, related to its goal-directed character, attentional suppression can be applied flexibly to the internal representation of any object or signal that interferes with the coherent performance of a task. The case for flexibility in the inhibitory component of attention is nicely illustrated in work by Tipper (1985) in the context of his negative priming paradigm. On a typical trial, subjects view a stimulus, such as a drawing or a letter, colored in red ink, superimposed over a second, to-be-ignored stimulus, colored in green ink. Subjects are instructed to produce the name associated with the stimulus in red ink, and to ignore the stimulus in green ink, supposedly included to make the task more difficult. If the red item on the subsequent trial is the to-be-ignored stimulus from the previous trial, subjects are slower to provide the appropriate name than if an irrelevant stimulus had been ignored on the previous trial—a phenomenon known as negative priming. Interestingly, negative priming does not depend on the existence of a priori semantic relationships between the attended and ignored items: attending to a saxophone while ignoring a picture of a dog results in substantial negative priming, despite their dissimilarity. The effect suggests a general inhibitory mechanism that can be directed at any representation impeding the coherent production of a response, regardless of previously established positive or negative associations in memory. Such an inhibitory mechanism might be regarded as a basic component of executive control of thought and action (Logan, 1985; Logan & Cowan, 1984). Indeed, Tipper has argued that, depending on the goals of the task, inhibitory mechanisms can be directed to different points in the information processing chain, such as at the stage of semantic representation or response production (Tipper, 1992). If inhibitory mechanisms in selective attention are truly this flexible, perhaps the same applies for retrieval inhibition. Such flexibility in the direction of inhibition seems at odds with the lateral inhibitory approach, insofar as we take lateral inhibition as an automatic process operating among items with established similarity relations.[4]

If a flexibly directed inhibitory mechanism causes retrieval inhibition, then any item interfering with the production of a memory target ought to be subject to inhibition, regardless of whether the interfering item is similar or shares a common retrieval cue with the target. An example of this flexibility may be the retroactive interference that occurs in the version of the paired associate paradigm in which stimuli are not shared across the two lists. Most of the factors proposed to underlie retroactive interference (e.g., response competition, unlearning, lateral inhibition) predict the greatest interference between lists with identical stimulus terms (Dog–Rock, Dog–Sky), with no interference predicted to occur between lists with completely distinct stimulus terms (Dog–Rock, Ball–Sky). Yet, although lists with shared stimuli display the greatest interference, lists with differing stimuli also show interference. This finding has been used to argue for a general response-set suppression mechanism (Postman, Stark, & Fraser, 1968), the characteristics of which resemble attentional suppression mechanisms. According to Postman et al., retroactive interference results from a mechanism of response selection that suppresses the entire class of first-list responses as a whole, rather than individual stimulus–response associations. This process, occurring during the acquisition of second-list responses, involves a response selector mechanism that activates newly prescribed responses and inhibits earlier ones. Thus, Postman et al.'s response selector mechanism performs the functions that attentional mechanisms purportedly achieve in selective attention tasks: the activation of the target representation and the inhibition of activated competitors. Though not expressed in attentional terms, Postman et al.'s proposal of a response-set selection mechanism can be regarded as an early instance of an attentional suppression theory in the domain of memory.

Perhaps more direct evidence for a flexibly controlled attentional suppression mechanism producing retrieval inhibition in long-term memory comes from work on directed forgetting (e.g., Bjork, 1972; Coe, Basden, Basden, & Fikes, 1989; Epstein, 1972; Geiselman & Bagheri, 1985; Roediger & Crowder, 1972). In the typical directed forgetting paradigm, subjects learn a list of 10 unrelated nouns, after which they are either told to forget that list (the "forget" group) because the wrong words had been given to them, or to remember that list (the "remember" group). Both groups then receive a second list of nouns. After a brief retention interval, both groups are tested on their memory for both lists. Typically, forget subjects' recall of the first list is impaired relative to remember subjects' recall of that list, whereas forget subjects' recall of the second list is facilitated. Such impairment occurs even to incidentally encoded items on the first list, indicating that the effect does not merely reflect a deficit in rehearsal on the first list for forget subjects (Geiselman, Bjork, & Fishman, 1983). As in the negative priming paradigm discussed previously, there need exist no a priori semantic relationships between items on the first list and items on the second list for impairment to occur. Because there are no preexisting competitive relationships among items, and because impairment only occurs when subjects are directed to forget the earlier list, it is difficult to see how automatic lateral inhibitory mechanisms could be invoked to account for this effect. Bjork (1989) argues that directed forgetting may arise from the action of a suppression mechanism that subjects can actively direct at memory representations to alleviate effects of proactive interference. This interpretation is supported by recent work showing that directed forgetting effects are greatly attenuated in elderly populations thought to have a deficit in attentional inhibition (Radvansky, Zacks, & Hasher, 1991).

Although treating retrieval as the internal focus of attention may be interesting, and is likely to be a fruitful source of empirical predictions, it is not clear that all forms of retrieval inhibition can be accounted for within this framework. In particular, directed inhibitory mechanisms seem most applicable in cases in which there are right and wrong responses to be given on retrieval attempts—as is the case in our retrieval-

---

[4] This argument makes it unclear whether Carr and Dagenbach's (1990) center-surround theory of semantic retrieval is a pure attentional suppression mechanism. Carr and Dagenbach speculate that their center-surround inhibitory mechanism may employ lateral inhibitory mechanisms such as those used in Walley and Weiden's (1973) theory of attention.

practice paradigm and in the retroactive interference procedure. In our retrieval-practice paradigm, the category-stem retrieval cues require subjects to retrieve particular exemplars; activation of competitors is likely to impede report of the appropriate item. In retroactive interference, study-test trials on second-list items require subjects to give only the most recent response, rendering first-list items as errors. In part–set cueing inhibition, however, subjects may report any item at any time. Activation of competing items is not inappropriate, because the report of any is a correct response. It is not clear, then, why retrieval inhibition should be necessary in this task.

IV. A CRITERION FOR INHIBITION IN LONG-TERM MEMORY

Most of the "inhibitory" phenomena reviewed in the introduction can be caused by the variety of associative and cue bias mechanisms outlined in the noninhibitory section of our taxonomy. Although each of these processes, from the standpoint of someone researching the causes of retrieval failure, is a potentially interesting source of forgetting (deserving investigation in its own right), these alternative mechanisms pose a problem for anyone interested in exploring the properties of inhibitory processes: How can genuine inhibition be differentiated from occlusion, resource diffusion, associative decrement, or biases in retrieval cues? For any inhibitory paradigm, some unknown mixture of these mechanisms may contribute to impairment. This ambiguity makes it difficult not only to assess the quantitative and qualitative properties of inhibition, but also to establish the contribution of inhibitory mechanisms at all. Criteria for what constitutes evidence for an inhibitory mechanism are clearly needed.

The previous section reviewed several theoretical mechanisms that might produce genuine inhibition of memory items. When we say an item is inhibited in the theoretical sense, we assert that the representation of that item has suffered a decrease in its activation. The difficulty in believing that the performance impairments reviewed in the introduction reflect inhibition arises because it is never clear that anything has truly happened to the representations of the nonretrieved items. In this section, we review a minimal criterion (i.e., one that is necessary, but may not be sufficient) that a phenomenon must meet to be convincingly established as inhibitory—a criterion referred to as *cue-independent impairment.* We then review the independent probe method, a technique developed by Anderson and Spellman (1991a, 1991b, 1993) for establishing cue-independent forgetting. We present evidence using this paradigm that despite our skepticism has led us to believe that inhibitory mechanisms do produce forgetting in long-term memory.

Anderson and Spellman (1991a, 1991b, 1993) noted that an appropriate test for inhibition must measure changes in the target representation separately from changes to associations and cues. Although associative and cue bias factors seem hopelessly confounded with target inhibition in current paradigms, these factors may be separated if we consider the implication of one crucial observation: Whereas inhibitory models localize impairment to the item itself, factors such as associative bias and cue bias are extrinsic to the target. The fact that extrinsic factors cause impairment implies that changes in the retrieval cue by which an item is accessed should eliminate impairment arising from noninhibitory mechanisms. Any remaining impairment, therefore, should reflect changes in the state of the target, presumably caused by retrieval-based inhibitory processes. For example, if retrieval practice of Fruit–Orange truly inhibits the representation of Banana, recall of Banana should decrease regardless of whether subjects are cued with the category Fruit or with an independent cue such as Monkey. Thus, associative bias and cue bias models predict cue-dependent forgetting (Tulving, 1974), whereas target bias (inhibitory) models predict that forgetting should be independent of the cue used to test the impaired item. If we observe impairment when recall is tested with a cue that is unrelated to the strengthened item, we have a measure of inhibition free from the influences of the noninhibitory mechanisms outlined in this chapter.

To determine whether impairment caused by an inhibitory paradigm is cue independent, independent probes must be used to test recall of impaired items—that is, recall must be tested with probes sufficiently related to impaired items to be effective cues, but that are both (1) unrelated to strengthened competitors and (2) unrelated to the cue shared by the impaired item and its strengthened competitor. For instance, in the preceding example (cuing Banana with Monkey, given that Fruit–Orange was practiced), Monkey is related to Banana, but not to Orange, so it is likely to activate Banana selectively. Because Monkey is not strongly related to Fruit, the common cue linking Orange and Banana, its presentation will not activate Fruit, avoiding the indirect activation of Orange. Thus, Monkey serves as an independent probe for Banana. Following Anderson and Spellman (1991a, 1991b, 1993), we will refer to the use of such an independent cue to assess the activational state of an item as the *independent probe method.*

Anderson and Spellman (1991a, 1991b, 1993) applied the independent probe method to determine whether retrieval-induced forgetting (Anderson, Bjork, & Bjork, in press) is caused by occlusion or by genuine inhibitory processes. They reasoned that if impairment in the retrieval-practice paradigm resulted from lateral inhibitory processes among category exemplars, impairment should be cue independent. To test this prediction, Anderson and Spellman used Anderson et al.'s procedure, but with one crucial change: Unlike Anderson et al.'s work, in which categories were as unrelated as possible, Anderson and Spellman designed their materials so that items from some pairs of categories were similar to one another. Specifically, when two categories were members of the *related category condition*, three members of each category were unique to that category and three members could be categorized under the paired category. Figure 11 illustrates their stimulus materials (with only two items per category, for simplicity). In the example illustrated in Figure 11, the items Blood and Bread are unique to their respective categories, Red and Food, in the sense that they are not categorizable under each other's category label (e.g., Blood is not usually considered a Food). In contrast, Tomato and Radish, though appearing under the categories Red and Food, respectively, within the context of the experiment, may equally well be categorized under each other's experimental category (e.g., Tomatoes are also Food); that is, strong a priori associations exist between these exemplars and another category name (depicted by dotted lines). Anderson and Spellman used these materials to distinguish occlusion from genuine inhibition in the following manner. If Red–Blood is practiced, then, according to lateral inhibition, Tomato should suffer because of its membership in the Red category. Because Radish is also subsumed under the Red category (though only in semantic memory), it should also be impaired by practicing Red–Blood, even though Radish was studied and tested under the Food category (an independent probe). If occlusion causes retrieval-induced forgetting, studying and testing Radish under the Food category should provide a retrieval route to Radish uninfluenced by the practice of Red–Blood; Blood is not usually Food, so Blood should not intrude given the cue Food. Thus, occlusion predicts that practicing Red–Blood should impair Red–Tomato, but not Food–Radish.

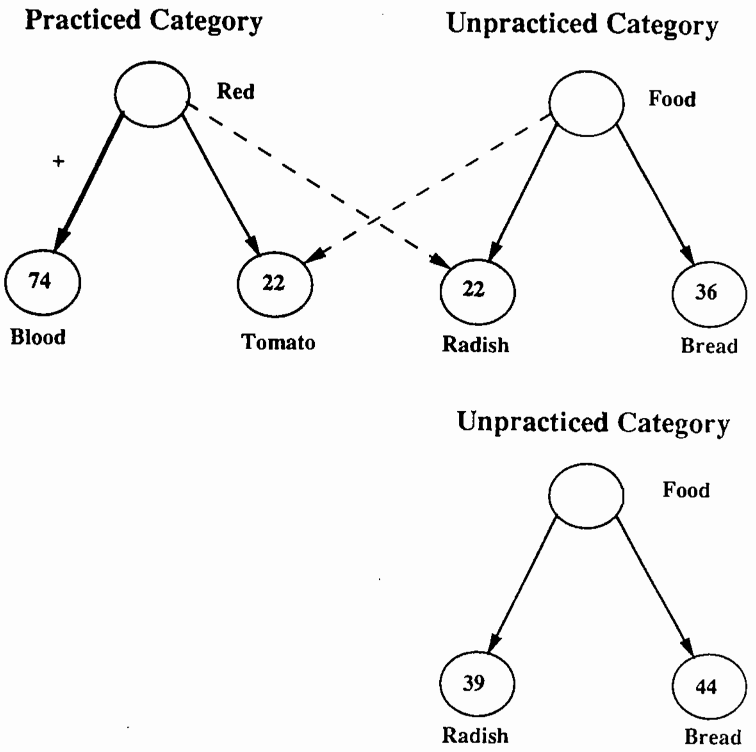

FIGURE 11

Design and results from a study by Anderson and Spellman (1991a, 1991b, 1993). The top materials illustrate the related-category condition, in which two categories with semantic overlap (one category—Red—partially practiced, the other unpracticed) are studied. The bottom illustrates the unrelated-category condition in which only one of the above paired categories is studied; it serves as the control against which to measure inhibition in the related-category condition. Recall percentages for each condition appear in the nodes for the relevant item.

Figure 11 displays results from a representative study by Anderson and Spellman. Mean recall scores are aligned with example items from the various experimental conditions. Replicating Anderson, Bjork, and Bjork (in press), retrieval practice significantly improved performance on unique items (e.g., Red–Blood) relative to the comparable baseline items (e.g., Food–Bread), but only at the expense of performance on other items in the practiced category (e.g., Red–Tomato). Of greater interest for present purposes is the impaired performance on Red items studied and tested under the Food category (e.g., Food–Radish), relative to those same items when Red things were not studied by subjects. Interestingly, impairment on these cross-categorizable items does not seem lessened by their encoding under an independent category context. This basic finding was replicated several times using different materials. Thus, Anderson and Spellman's results support the notion that retrieval practice impairs nonpracticed items via active inhibitory processes, not because of associative or cue bias mechanisms. Impairment can occur even when memory is tested with an independent retrieval cue.[5]

Models other than occlusion can also be ruled out with this demonstration of cue-independent impairment—at least with respect to cross-

---

[5] It should also be noted that the remaining items such as Food–Bread, which are not categorizable under the practiced category, Red, were also impaired (8%) by practicing Red–Blood, but not as much as the critical items (17%). Because Bread cannot fall under the Red category, its impairment appears to be inconsistent with a lateral inhibitory model. Subsequent experiments by Anderson and Spellman replicated this unusual finding and established the conditions necessary for it to occur. Careful analysis reveals that this finding is indeed predicted by lateral inhibitory models as well as by certain distributed models of retrieval inhibition (Anderson & Spellman, 1993).

category inhibition. To see this, consider what models, such as resource diffusion, associative decrement, meaning bias, and context bias, would predict in the independent probe paradigm. According to resource diffusion, practicing Red–Blood should decrease the flow of activation for Red–Tomato, because Blood and Tomato compete for activational resources from a common cue. Furthermore, such practice might decrease the flow of activation across the semantic link connecting Red and Radish, even though Radish was studied as a Food. Because Radish was tested using the Food cue, however, decreased efficacy of the Red–Radish semantic link is irrelevant to recall performance. Also, because the rate at which activation spreads across the Food–Radish link should be unaffected by practice of Red–Blood, performance on Radish should be unaffected. Similar logic applies to associative decrement and meaning bias models. As long as performance is tested with an independent cue, such as Food, bias of the Red cue, or associative decrement of the Red–Radish semantic link should be irrelevant. Finally, because Food was an unpracticed category, subjects ought to use the original learning context to search memory; thus, performance on Food–Radish should be unimpaired, according to context bias models.

This example illustrates the usefulness of the cue independence criterion for establishing inhibition. Because inhibitory mechanisms focus on the state of the item itself, rather than on the associations or the cue, they predict impairment where there should be none, according to noninhibitory models. Furthermore, there is nothing in the logic of cue-independent impairment that restricts the criterion to episodic memory recall; thus it should apply to inhibitory phenomena across a variety of domains. If a given phenomenon can be shown to meet this criterion, then, in most cases, all models on the left branch of the taxonomy in Figure 1 can be ruled out as necessary causes of the impairment. It should be emphasized, however, that we believe cue independent to be a minimal criterion for the establishment of a phenomenon as inhibitory. Further elaboration of the properties of items demonstrating cue-independent impairment may render the inhibitory interpretation unsatisfactory. However, given the inadequacy of the noninhibitory alternatives at explaining the results of Anderson and Spellman, and the plausibility of the functional role of inhibitory processes in retrieval, we interpret the results of Anderson and Spellman as evidence for the existence of inhibitory mechanisms in long-term memory.

V. THREE CHALLENGES FOR A THEORY OF RETRIEVAL INHIBITION

The previous sections emphasized the variety of mechanisms that may contribute to so-called inhibitory effects in long-term memory. Clearly, empirical work must be done to clarify the relative contributions of such mechanisms in inhibitory paradigms. As we have indicated, however, we believe there is good reason to think that true inhibitory mechanisms at least play a role in some cases of retrieval inhibition—particularly where impairment is demonstrated to be cue independent. But even if impairment in a variety of retrieval inhibition paradigms is shown to be cue independent, there remain many theoretical issues regarding the specific nature of this inhibition. In this section, we present three theoretical challenges for those interested in understanding retrieval inhibition and its function in long-term memory.

A. Explain Why Retrieval Inhibition Lasts So Long

The notion of a "true" inhibitory process underlying retrieval inhibition rests on a theoretical comparison to inhibition on the neuronal level. Frequently, it is unclear whether this comparison implies that retrieval inhibition genuinely reflects the consequences of neuronal inhibition, or whether inhibition is to be taken metaphorically. Although it is clear that retrieval inhibition does not represent hyperpolarization of individual neurons, as connoted in the simple network diagrams used throughout the chapter, it can be argued that it reflects inhibition of a population of neurons, the relationships among which represent the item. If the aggregate behavior of this system of neurons reflects that of its component neurons, retrieval inhibition may behave like neuronal inhibition. If retrieval inhibition is shown to behave similarly to neuronal inhibition, a neuronal interpretation of the phenomenon will become more convincing.

One characteristic of retrieval inhibition that seems at odds with a neuronal interpretation is the difference in the durability between the two phenomena. Whereas the result of an inhibitory impulse into a neuron is quite brief (between a few milliseconds and a few seconds; see Kandel & Schwartz, 1991), retrieval inhibition appears to last far longer. For example, in a classic series of studies, Postman, Stark, and Fraser (1968), demonstrated that retroactive interference lasts as long as 26 min. In our retrieval-induced forgetting paradigm, we still observed impairment after a 20-min retention interval (Anderson & Bjork, 1990; Anderson et al., in press; Anderson & Spellman, 1991a, 1991b, 1993). How can retrieval inhibition last so long if it is caused by neuronal inhibitory mechanisms? Although many researchers consider inhibitory mechanisms a viable interpretation of retrieval inhibition, little attention has been devoted to this glaring inconsistency with the properties of neuronal inhibition.

Although we will not attempt to answer this question here, we will suggest three alternatives that can explain why retrieval inhibition endures as long as it does. The first interpretation is that it does not truly endure throughout the retention interval; rather, it may be the dominance of strengthened exemplars over nonstrengthened competitors that endures. An enduring facilitation of strengthened exemplars may, at test, recreate inhibitory effects on nonstrengthened items because of the propensity for strengthened items to be reported first. Thus, the apparent discrepancy in the durability of inhibitory effects may not truly exist. Results from a study by Anderson et al. (in press), however, argue that such output interference is not the sole source of inhibitory effects. Anderson et al., employing the retrieval-practice procedure reviewed earlier, controlled the order in which subjects reported category exemplars by substituting a category-plus-stem completion test for the normal category-cued free recall procedure. By testing practiced items after unpracticed items in a category, Anderson et al. removed output interference as a potential cause of inhibition. Inhibition on unpracticed items was still observed, though decreased in magnitude relative to the condition in which practiced items were reported first. These findings were replicated using a somewhat different paradigm (Anderson & Spellman, 1991a, 1991b, 1993), reinforcing the interpretation that inhibitory effects are, in fact, quite durable. Thus, although persisting dominance of strengthened items may produce new inhibitory effects on each subsequent test, they appear only to compound an existing inhibitory effect.

A second approach to explaining the durability of retrieval inhibition is to recognize that such inhibitory effects represent the aggregate behavior of complex systems of neurons. Although the behavior of these systems might reflect that of its component neurons, it seems at least plausible that the behavior of the complex system will vary in quantitative if not qualitative respects from the behavior of its constituents. Thus, inhibition of the higher level system may behave like inhibition on the neuronal level, but it may be manifested on a larger time scale. Though underspecified, this class of possibilities cannot truly be ruled out in the absence of a characterization of the relevant neuronal circuitry.

Finally, inhibitory processes may be actively sustained throughout the retention interval. Recall performance should be impaired if inhibition is continuously applied to the representation. Sustained inhibition might be mediated by the sustained preparation of practiced responses during the retention interval via lateral inhibitory links, perhaps to prevent rebound effects from disrupting performance. However, it is not clear why sustained preparation should be necessary in many of the tasks in which enduring inhibition is observed. For example, in our retrieval-practice paradigm, subjects perceive the retrieval-practice phase as the final test; thus, they should not be actively maintaining practiced responses in a state of preparedness throughout the retention interval. Any such theory of the durability of inhibitory effects should be accompanied by a theory of response preparedness.

B. Specify the Relationship between Attentional Inhibition and Retrieval Inhibition

As alluded to in our discussion of inhibitory mechanisms, retrieval inhibition may bear an interesting relationship to phenomena in selective attention. At a minimum, attentional suppression and retrieval inhibition appear functionally analogous: tasks in both domains require the isolation of a target representation from among a set of active alternatives to produce a response. Even granting this functional analogy, however, the specific relation between these phenomena remains unclear. We believe that clarification of the relation of these two areas is theoretically important and would yield interesting empirical ideas. We next consider several interpretations of the relationship between these phenomena.

First, the computational problems in selective attention and memory retrieval may not merely be similar—they may be exactly the same problems solved with the same attentional focusing mechanisms. Focusing attention to a target item may, in some circumstances, result in long-term suppression of distractors similar to what we observe in retrieval inhibition paradigms. A common inhibitory mechanism underlying both retrieval inhibition and attentional suppression does not seem far-fetched when we step away from the traditional divisions made between these domains. For example, the color–word responses inhibited in the Stroop version of the distractor suppression effect (Neill, 1977) are clearly stored in memory—thus, impairment in naming the color on subsequent trials might just as well be referred to as retrieval inhibition. Similarly, the interfering quality of distractors in selective attention arises from their activation of memory representations that disrupt focusing processes. It seems a relatively simple extension to allow activation of distractors to arise internally through spreading activation from other memory representations. Thus, resolving competition during memory retrieval might be regarded as selectively attending to memory items. Although memory researchers often focus on the process of recall for its own sake, it is perhaps more general and realistic to consider memory as having evolved in service of action and perception. When considered in this light, a common mechanism underlying these phenomena should not seem surprising.

A second possibility is that attentional suppression and retrieval inhibition may be produced by similar mechanisms at different levels in the nervous system. To understand this possibility, consider an example of inhibition in motor control. When the flexor muscle of a limb receives an impulse, inhibitory signals are normally sent to the corresponding extensor muscles, preventing them from contracting during flexor contraction. The reverse occurs when the extensor is contracted. Mutual inhibition prevents the paralysis that might result if the opposing muscles were employed in parallel for physically conflicting actions. This mechanism might be functionally similar to that occurring in retrieval inhibition; in retrieval practice, for example, the task requires that one response be given, although many incompatible responses become activated. Because one response must be produced, others might be inhibited. Even the most steadfast supporter of a comparison between motor and retrieval inhibition, however, would never suggest that the same specific circuitry (i.e., circuitry at the same physical location in the nervous system) underlies both processes. At best, these might be functionally analogous mechanisms at two levels of organization in the nervous system—that is, they may both be instances of the way the nervous system solves the problem of selection (see Allport, 1987; Neumann, 1987; Tipper, 1985, for discussions of the role of inhibition in selection). Similarly, although attentional suppression and retrieval inhibition may involve similar inhibitory circuitry, they may occur in functionally distinct systems, or at different levels of organization in the nervous system.

Third, it is possible for attentional suppression and retrieval inhibition to be functionally analogous, yet mechanistically distinct. This would be especially plausible given the discovery of both qualitative and quantitative differences between the inhibitory phenomena. However, differences in implementation should not necessarily lead to the abandonment of the functional analogy; it may still be appropriate to view the two mechanisms as different ways of achieving the selection of a single response from among a set of distracting competitors.

Finally, attentional suppression and retrieval inhibition may be functionally distinct, but mechanistically analogous. We include this possibility because we wish to emphasize that all cases of inhibition need not reflect functionally similar circumstances. The assumption of functional similarity between retrieval inhibition and attentional suppression seems plausible because impairment in the two cases appears precipitated by circumstances in which a response selection must be made. It is possible, however, that similar inhibitory circuitry might serve different functions across the domains of selective attention and memory retrieval. For example, one can acknowledge that selection of the appropriate memory trace is a necessary component of retrieval without accepting the premise that inhibitory mechanisms subserve that function. Impairment of related memory traces during selection may represent the consequence of some functional demand that is merely correlated with selection of the memory trace. Though this logical possibility should be considered, its advocates need to (1) specify a theory of the function of inhibition in memory retrieval that is not based on selection, and (2) specify how selection in memory retrieval, unlike selection in attention, can occur without the aid of inhibition.

C. Provide a Computational Mechanism That Can Capture the Flexibility of Inhibition

In our discussions of the relationship between attentional suppression and retrieval inhibition, we speculated that both phenomena may be produced by a flexibly controlled inhibitory mechanism. The active character of this hypothetical process is motivated by two aspects of attentional suppression. First, effects such as negative priming occur when subjects attempt to ignore a distracting stimulus while attending to a target. Subjects know they must ignore the distracting image and seem able to wield control over the process. Second, it seems implausible that an automatic lateral inhibitory process underlies negative priming because such processes are usually thought to occur between computational units with some a priori competitive relationship (e.g., logical or physical incompatibility of responses, similarity). That negative priming occurs to the semantic representation of a distracting item that is semantically unrelated to the target (e.g., a saxophone and a dog) leads one to question whether lateral inhibition is a good model of negative priming. Thus, it seems plausible that negative priming results from a controlled, goal-directed inhibitory process.

These speculations about a controllable inhibitory process, though intriguing, are relatively vague about the mechanism's functional characteristics and how they might be implemented on a computational level. Exactly how flexible might such a mechanism be? Could any representation be inhibited at any time, or must that representation be actively interfering with the processing of some other stimulus or response? If this process can inhibit any representation, how is this possible on a mechanistic level? Normally, inhibition is thought to spread from one computational unit across inhibitory connections to the to-be-inhibited unit. Must we assume that there exists some unit or inhibitory circuit that has inhibitory links to all possible representations? How would such a structure inhibit items selectively in coordination with patterns of controlled activation? Might such focusing of attention on internal representations operate by principles similar to those employed by thalamic gating mechanisms thought to underlie visual selective attention (Crick, 1984; Laberge, 1990; Scheibel, 1981; Yingling & Skinner, 1977)? If we are to make empirical and theoretical progress in differentiating controlled from automatic forms of inhibition, a clearer theory of how controlled inhibition occurs must be developed.

VI. SUMMARY AND CONCLUSIONS

A central mission of this volume has been to examine the extent to which empirical evidence justifies the postulation of inhibitory mechanisms in the variety of domains in which they have become popular: selective attention, memory retrieval, and language processing. This question arises because, for many empirical phenomena, theories proposing an excess of facilitation on opposing responses can account for performance impairments as effectively as inhibitory models. Such ambiguity renders use of inhibitory constructs a matter of theoretical taste. The present chapter discusses an instance of this issue arising in research on forgetting from long-term memory: the issue of whether retrieval inhibition—impaired accessibility to demonstratably available memory items—is produced by inhibitory or noninhibitory mechanisms. We addressed this issue through two contributions: (1) a review and integration (in the form of a taxonomy) of the noninhibitory and inhibitory mechanisms that might underlie retrieval inhibition, and (2) an empirical criterion by which these models might be distinguished.

In our view, noninhibitory models are those models that do not appeal to a subtractive process reducing the level of activation of impaired items. Noninhibitory models of retrieval inhibition can be grouped into three general classes according to the source of bias in performance. Associative bias models emphasize the role of associations linking the retrieval cue and the target memory; cue bias models attribute retrieval inhibition to biases in the meaning of the functional retrieval cue toward features of frequently used items; and executive control models assert that impaired performance arises from biases in executive control factors such as the point of search termination and the familiarity threshold for reporting items. Much of the data on retrieval inhibition are consistent with at least some of these noninhibitory mechanisms, making them difficult to distinguish from inhibitory models, and, indeed, from each other.

Inhibitory models are those models that attribute performance impairment to a decrease in the activation of the affected item as a result of an activation-reducing mechanism. At least two general classes of inhibitory models may be distinguished: lateral inhibitory and attentional suppression models. Lateral inhibitory models normally presuppose a priori competitive relationships among items bearing certain characteristics (e.g., similarity, logical incompatibility)—competition that is built into the structure of memory in the form of inhibitory connections linking incompatible items. Such competitive processes are often thought to be automatic. Attentional suppression models, on the other hand, propose flexible, goal-directed inhibitory processes that can be applied to any representation whose activation interferes with ongoing processing. Though less developed computationally, such models suggest an interesting link between inhibitory phenomena in selective attention and long-term memory.

From the standpoint of theoretical parsimony, it seems unnecessarily complex to postulate inhibitory processes such as those in lateral inhibition and attentional suppression to account for data that can be explained adequately by ostensibly simpler, noninhibitory models. However, as stressed in our discussion of lateral inhibitory models, many noninhibitory models propose additional processes, that although not labeled as inhibitory, may disguise or rename an inhibitory process. For example, both occlusion (e.g., McGeoch, 1942; Raajimakers & Shiffrin, 1981; Rundus, 1973) and resource diffusion (e.g., Anderson, 1976, 1983) models, as summarized earlier, make limited capacity assumptions with respect to an output channel and activational resources, respectively. Yet, it is never explained how those capacities come to be limited (see also, Neumann, 1987, for a related critique of limited resource theories of attention in general). In our section on lateral inhibitory models, we developed lateral inhibitory circuits illustrating how the behavior of such noninhibitory processes might be modeled via inhibitory mechanisms. To the extent that limited-capacity assumptions are necessary to implement noninhibitory mechanisms, and to the extent that those assumptions bury inhibitory processes, noninhibitory models may not be more parsimonious.

But theoretical parsimony is not the only basis on which to prefer one class of models over another. We argue that inhibitory processes may be distinguished from noninhibitory alternatives by a general criterion developed by Anderson and Spellman (1991a, 1991b, 1993): cue-independent impairment. Applied to episodic recall paradigms, cue-independent impairment requires that recall deficits of an item be measurable from a separate retrieval cue (i.e., separate from the one associated to both the strengthened competitor and the target), something which should occur if the representation of the item is truly less active. We review a study by Anderson and Spellman demonstrating cue-independent impairment. This finding cannot be explained by any of the noninhibitory models in our taxonomy and suggests that inhibitory mechanisms play an important role in retrieval-induced forgetting. Many challenges remain, however, in the theoretical characterization of this inhibition (e.g., how does it last so long?) and how it might relate to other inhibitory phenomena such as attentional inhibition.

Although retrieval inhibition should generally be regarded as a descriptive label, we believe that evidence of cue-independent impairment warrants the use of the term in its stronger mechanistic sense—at least in the case of retrieval-induced forgetting. The functional conditions under which this impairment occurs mirror those present in attentional suppression, suggesting an exciting computational principle—and perhaps a common mechanism—mediating the phenomena.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by Contract MDA 903-89-K-0179 from the Army Research Institute to Keith J. Holyoak. The authors gratefully acknowledge John Hummel and Tom Wickens for extensive comments on early drafts of this chapter, and Laura Da Costa for editorial assistance on the final version. We also wish to thank Dale Dagenbach and Tom Carr for insightful comments and recommendations.

REFERENCES

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–420). Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6*, 451–474.

Anderson, J. R. (1976). *Language, memory and thought.* Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, M. C., & Bjork, R. A. (1990, November). *Category-specific retrieval inhibition.* Paper presented at the annual conference of the Psychonomic Society, New Orleans, LA.

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (in press). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Anderson, M. C., & Spellman, B. A. (1991a, June). *Retrieval practice inhibits similar memories, regardless of whether common cues are shared.* Poster presented at the annual conference of the American Psychological Society, Washington, D.C.

Anderson, M. C., & Spellman, B. A. (1991b, July). *Does genuine inhibition occur in long-term memory?* Paper presented at the International Conference on Memory, Lancaster, England.

Anderson, M. C., & Spellman, B. A. (1993). *On the status of inhibitory mechanisms in cognition: Memory Retrieval as a Model Case.* Manuscript submitted for publication.

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58*, 95–105.

Barsalou, L. W. (1982). Context dependent and context independent information in concepts. *Memory & Cognition, 10*, 82–93.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101–140). Cambridge: Cambridge University Press.

Barsalou, L. W. (1989). Intraconcept and interconcept similarity. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning.* New York: Cambridge University Press.

Basden, D. R., Basden, B. H., & Galloway, B. C. (1977). Inhibition with part-list cuing: Some tests of the item strength hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 100–108.

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

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington, DC: Winston & Sons.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Lawrence Erlbaum.

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of a category-specific inhibition. *Memory & Cognition, 11*, 500–510.

Brown, A. S. (1991). The tip of the tongue experience: A review and evaluation. *Psychological Bulletin, 109*, 204–223.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30*, 542–579.

Butter, C. M. (1969). Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in *Macaca Mulatta. Physiology and Behavior, 4*, 163–171.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 341–350.

Chandler, C. C. (1989). Specific retroactive interference in modified recognition tests: Evidence for an unknown cause of interference. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*, 256–265.

Chandler, C. C. (1993). Accessing related events increases retroactive interference in a matching recognition test. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 967–974.

Coe, W. C., Basden, B. H., Basden, D., & Fikes, T. (1989). Directed forgetting and posthypnotic amnesia: Information processing and social contexts. *Journal of Personality and Social Psychology, 56*, 189–198.

Crick, F. (1984). The function of the thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences of the U.S.A., 81*, 4586–4590.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Lawrence Erlbaum.

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 and Cognition, 16*, 328–340.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension of intelligence. *Intelligence, 15*, 157–173.

Epstein, W. (1972). Mechanisms in directed forgetting. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 147–192). New York: Academic Press.

Erdelyi, M. H., & Goldberg, B. (1979). Let's not sweep repression under the rug: Toward a cognitive psychology of repression. In J. F. Kihlstrom & F. J. Evans (Eds.), *Functional disorders of memory* (pp. 355–402). Hillsdale, NJ: Lawrence Erlbaum.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62*, 369–377.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–191). New York: Wiley.

Freud, S. A. (1952). *A general introduction to psychoanalysis* (J. Riviere, trans.). New York: Washington Square Press. (Originally published, 1920.)

Fuster, J. M. (1989). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe.* New York: Raven Press.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory & Cognition, 13*, 51–62.

Geiselman, R. E., Bjork, R. A., & Fishman, D. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112*, 58–72.

Gernsbacher, M. A., Barner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 430–445.

Hartman, M., & Hasher, L. (1991). Aging and suppression: Memory for previously relevant information. *Psychology and Aging, 6*, 587–594.

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.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York: Academic Press.

Hinton, G. E., & Plaut, D. C. (1987). Using fast weights to deblur old memories. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society* (pp. 177–186). Hillsdale, NJ: Lawrence Erlbaum.

Hintzman, D. L. (1978). *The psychology of learning and memory* (pp. 297–304). San Francisco: W. H. Freeman.

Kandel, E. P., & Schwartz, J. H. (1991). *Principles of neural science.* New York: Elsevier.

Katz, J. J., & Postal, P. (1964). *An integrated theory of linguistic descriptions.* Cambridge, MA: MIT Press.

Keele, S. W., & Neill, W. T. (1978). Mechanisms of attention. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception IX.* (pp. 3–47). New York: Academic Press.

Kihlstrom, J. F. (1983). Instructed forgetting: Hypnotic and nonhypnotic. *Journal of Experimental Psychology: General, 112*, 73–79.

Laberge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience, 2*, 358–372.

Levy, W. B., & Steward, O. (1983). Temporal contiguity requirement for long-term associative potentiation/depression in the hippocampus. *Neuroscience, 8*, 791–797.

Lewandowsky, S. (1991). Gradual unlearning and catastrophic interference: A comparison of distributed architectures. In W. E. Hockey and S. Lewandowsky (Eds.), *Relating theory and data: Essays in honor of Bennet B. Murdock* (pp. 445–476). Hillsdale, NJ: Lawrence Erlbaum.

Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology, 7*, 311–335.

Lindsay, D. S. (1993). Eyewitness suggestibility. *Current Directions in Psychological Science, 2*, 86–88.

Loftus, E. F. (1979). The malleability of memory. *American Scientist, 67*, 312–320.

Loftus, E. F., & Loftus, G. R. (1980). On the permanence of stored information in the human brain. *American Psychologist, 35*, 409–420.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*, 295–327.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica, 60*, 193–210.

Luce, R. D. (1959). *Individual choice behavior.* New York: Wiley.

Luria, A. R. (1966). *Higher cortical functions in man.* London: Tavistock.

Martin, E. (1971). Verbal learning theory and independent retrieval phenomena. *Psychological Review, 78*, 314–332.

Martindale, C. (1981). *Cognition and consciousness.* Homewood, IL: Dorsey Press.

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

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 109–164). San Diego: Academic Press.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 1*, 1–16.

McGeoch, J. A. (1936). Studies in retroactive inhibition: VII. Retroactive inhibition as a function of the length and frequency of presentation of the interpolated lists. *Journal of Experimental Psychology, 19*, 674–693.

McGeoch, J. A. (1942). *The psychology of human learning.* New York: Longmans, Green.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 3*, 173–203.

Mensink, G. J. M., & Raajimakers, J. G. W. (1988). A model of interference and forgetting. *Psychological Review, 95*, 434–455.

Mishkin, M. (1964). Perseveration of central sets after frontal lesions in monkeys. In J. M. Warren & K. Akert (Eds.), *The frontal granular cortex and behavior.* New York: McGraw-Hill.

Murdock, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology, 60*, 222–234.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4*, 648–654.

Neely, J. H., Schmidt, S. R., & Roediger, H. L. (1983). Inhibition from related primes in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, 196–211.

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

Neill, W. T. (1989). Lexical ambiguity and context: An activation-suppression model. In D. S. Gorfein (Ed.), *Resolving semantic ambiguity* (pp. 63–83). New York: Springer-Verlag.

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. (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 & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361–394). Hillsdale, NJ: Lawrence Erlbaum.

Nickerson, R. S. (1984). Retrieval inhibition from part–set cuing: A persisting enigma in memory research. *Memory & Cognition, 12*, 531–552.

Nicoll, R. A., Kauer, J. A., & Malenka, R. C. (1988). The current excitement in long-term potentiation. *Neuron, 1*, 97–103.

Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwarts, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenium Press.

Postman, L. (1971). Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth and Schlosberg's: Experimental psychology* (3rd ed.). New York: Holt, Rinehart & Winston.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1*, 19–40.

Postman, L., & Stark, K. (1969). The role of response availability in transfer and interference. *Journal of Experimental Psychology, 79*, 168–177.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior, 7*, 672–694.

Raajimakers, J. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 940–953.

Radvansky, G. A., Zacks, R. T., & Hasher, L. (1991, May). *Directed forgetting and aging.* Paper presented at the meeting of the Midwestern Psychological Association, Chicago, IL.

Ratcliff, R. (1990). Connectionist models of recognition memory: Constraints imposed by learning and forgetting functions. *Psychological Review, 97*, 285–308.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 163–178.

Reason, J. T., & Lucas, D. (1984). Using cognitive diaries to investigate naturally occurring memory blocks. In J. E. Harris & P. E. Morris (Eds.), *Everyday memory actions and absent-mindedness* (pp. 53–70). London: Academic Press.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition, 2*, 261–269.

Roediger, H. L., & Crowder, R. G. (1972). Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology, 3*, 244–254.

Roediger, H. L., & Neely, J. H. (1982). Retrieval blocks in episodic and semantic memory. *Canadian Journal of Psychology, 36*, 213–242.

Roth, E. M., & Shoeben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15*, 346–378.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations.* Cambridge, MA: MIT Press.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6*, 1–36.

Rundus, D. (1973). Negative effects of using list items as retrieval cues. *Journal of Verbal Learning and Verbal Behavior, 12*, 43–50.

Sandson, J., & Albert, M. L. (1984). Varieties of perseveration. *Neuropsychologia, 22*(6), 715–732.

Scheibel, A. B. (1981). The problem of selective attention: A possible structural substrate. In O. Pompeiano & C. Marsen (Eds.), *Brain mechanisms and perceptual awareness.* New York: Raven Press.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge: Cambridge University Press.

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). The list-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 179–195.

Slamecka, N. J. (1975). Intralist cueing of recognition. *Journal of Verbal Learning and Verbal Behavior, 14*, 630–637.

Sloman, S. A., Bower, G. H., & Roher, D. (1991). Congruency effects in part–list cuing inhibition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*, 974–982.

Sloman, S. A., & Rumelhart, D. E. (1992). Reducing interference in distributed memories through episodic gating. In A. Healy, S. Kosslyn, and R. M. Shiffrin (Eds.), *From learning theory to connectionist theory: Essays in honor of William K. Estes* (pp. 227–248). Hillsdale, NJ: Lawrence Erlbaum.

Stanton, P. K., & Sejnowski, T. J. (1989). Associative long-term depression in the hippocampus induced by Hebbian covariance. *Nature, 339*, 215–218.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*, 135–170.

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

Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. *Current Directions in Psychological Science, 1*, 105–109.

Todres, A. K., & Watkins, M. J. (1981). A part-set cueing effect in recognition. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 91–99.

Tulving, E. (1974). Cue-dependent forgetting. *American Scientist, 62*, 74–82.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Walley, R. E., & Weiden, T. D. (1973). Lateral inhibition and cognitive masking: A neuropsychological theory of attention. *Psychological Review, 80*, 284–302.

Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science, 9*, 51–74.

Watkins, M. J. (1978). Engrams as cuegrams and forgetting as cue-overload: A cueing approach to the structure of memory. In C. R. Puff (Ed.), *The structure of memory* (pp. 347–372). New York: Academic Press.

Yaniv, I., & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential bases for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 187–205.

Yingling, C. D., & Skinner, J. E. (1977). Gating of thalamic input to cerebral cortex by nucleus reticularis thalami. In J. E. Desmedt (Ed.), *Attention, voluntary contraction and event-related cerebral potentials: Progress in clinical neurophysiology* (Vol. 1, pp. 70–96). Basel: Karger.