In H.L. Roediger & F.I.M. Craik (Eds.), Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving (pp.309-330). Hillside, NJ: Erlbaum.

16

Retrieval Inhibition as an Adaptive Mechanism in Human Memory

Robert A. Bjork
University of California, Los Angeles

It is argued herein that inhibition plays an important role in higher-order as well as lower-order cognitive processes. One such form of inhibition for which there is accumulating evidence is retrieval inhibition, characterized by a loss of access to certain items that are, in fact, stored in memory. Some of that evidence is summarized in this chapter, and the possible adaptive role played by retrieval inhibition in the updating of human memory, in the ability of higher-order units in memory to act as units, and in the long-term retention of order information is outlined.

Because the word *inhibition* is used in a number of ways in the literature—often simply as a descriptor for empirical effects that are the opposite of *facilitation*—it is important to emphasize that "inhibition" is used here in the strong sense, as in *suppression;* that is, as the opposite of *excitation.* The distinction is important because effects labelled as "inhibitory" are often the consequence of strengthening or activating incompatible or alternative responses, rather than the consequence of directly suppressing or inhibiting the response of interest. Such distinctions are discussed more fully in the final section of this essay.

INHIBITORY MECHANISMS AS EXPLANATORY CONSTRUCTS

In the broadest sense, we know that inhibitory processes are as important as excitatory processes in human information processing. At the neural level, inhibitory and excitatory processes work together to convey sensory information. In the ontogeny of brain development the later-developing (higher order) brain structures inhibit, via descending fibers, the earlier-developing (lower order) structures, which permits the organism to override simple reflexes, then fixed-action patterns, then instrumental behaviors, and, eventually, to operate on internalized representations.

In our theories of human memory, however, inhibitory processes have played little or no role. There are two reasons for that, in my opinion. First, notions of inhibition or suppression in human memory have an unappealing association to certain poorly understood clinical phenomena, such as repression. Second, the information-processing approach to the study of human memory, grounded as it is in the computer metaphor, leads us to think in terms of processes like storing, scanning, grouping, erasing, and so forth. Notions like inhibition, suppression, unlearning, and spontaneous recovery are not easily compatible with the computer metaphor. It is my belief, however, that inhibitory processes *do* play a major role in higher cognitive processes, such as memory.

It is interesting that Endel Tulving, the man in whose honor this essay is written, has had so little to say—in his written work—about inhibitory mechanisms (a clear exception being Tulving & Hastie, 1972, which deals with inhibition of the blocking type that is discussed later in this essay). For the last quarter of a century, Endel Tulving has been a key player and often the prime mover in many of the important developments in the study of human memory. The fact that he has written so little on inhibitory processes illustrates the minor role the inhibition construct has played in our theorizing. In presenting Endel Tulving as a kind of case study, however, I must add that he has, in his informal comments to myself and others over the years, expressed with his typical vigor the need for theories of memory to incorporate inhibitory mechanisms—because, he has argued, memory phenomena are ultimately the product of the living brain, not the product of an information-processing architecture analogous to that in the typical computer.

At the current stage of research on human memory, we are poorly served by the computer metaphor. Thirty years ago, as an alternative to the stimulus–response approach, the information-processing approach was invaluable. It led us to distinguish between control processes (software) and structural features (hardware), and it opened the black box to theoretical speculations about the number and nature of the processing stages inside. We have come to realize, however, that in virtually every important respect, the human information processor is functionally nothing like the standard digital computer. From the standpoint of a computer scientist, there appears to be considerable value in drawing upon modern neuroscience and cognitive science to reconfigure the processing architecture of the next generation of computers. From the standpoint of a cognitive psychologist, there is no remaining value in the standard computer metaphor.

Let me illustrate the foregoing points with an example from my own research. Some years back, we looked at the directed-forgetting manipulation in the context of Sternberg's (1966, 1969) memory-scanning paradigm (Bjork, Abramowitz, & Krantz, 1970). On each trial, subjects were presented a string of 0, 1, 2, 3, or 4 digits in one color (Color A) at a 0.67 sec rate before a switch to 1, 2, 3, or 4 digits presented in another color (Color B). A change in color served to cue subjects to forget the first set of digits (the TBF digits); the second set of digits defined the positive or to-be-remembered (TBR) set for that trial. At the end of the string of digits, a single-digit probe was presented in white, and the subjects were required to respond "yes" or "no" (with a button push) as quickly as possible depending on whether the probe was or was not in the TBR set of digits on that trial. Thus, there were two types of negative probes: those that were presented on that trial but in the TBF rather than in the TBR set (F-NOs), and those that were not presented at all on that trial (N-NOs). It should be noted that a digit was never repeated from the TBF or the TBR set in the same trial.

One further procedural detail is important. Subjects could not ignore the digits presented in the first color because on 33% of the trials there was no change of color, which meant that the TBR or positive search set was the Color A digits. On such trials, then, the procedure was the same as the standard Sternberg paradigm. Furthermore, because there was no way for the subject to know on any given trial whether there would be a color change until that change actually occurred, the subject had to try to remember Color A items on all trials.

This experiment produced stunning results. Those results were never published, however, because they seemed to defy interpretation. The principle questions of interest, in advance, were the following: (a) would YES reaction times depend on the size of the TBF set; (b) would reaction times for F-NO and N-NO responses differ; and (c) would F-NO reaction times differ as a function of the size of the TBF set?

The answer to the second question is shown in Fig. 16.1, where reaction time is plotted as a function of the size of the TBR set, averaged across the size of the TBF set. All three types of responses bear the standard linear relationship to the size of the TBR (positive) set, with F-NO responses being slower than N-NO responses by a relatively constant amount.

The answers to the other two questions are shown in Fig. 16.2, in which the reaction times are plotted as a function of the size of the TBF set, averaged over the size of the TBR set. None of the three response types vary as a function of the size of the TBF set.

At the time that we obtained these results, they defied description by any of the single-process (serial-parallel) search models then available. How could YES responses have shown no effect of the TBF set whatsoever when, at the same time, F-NO responses were slower than N-NO responses? If subjects could limit their search to the TBR set (as indicated by the YES results), why should it matter whether a negative probe was presented in the TBF set or not? And, as it did matter (F-NO responses were slower than N-NO responses), how can F-NO reaction times *not* depend on the size of the TBF set?

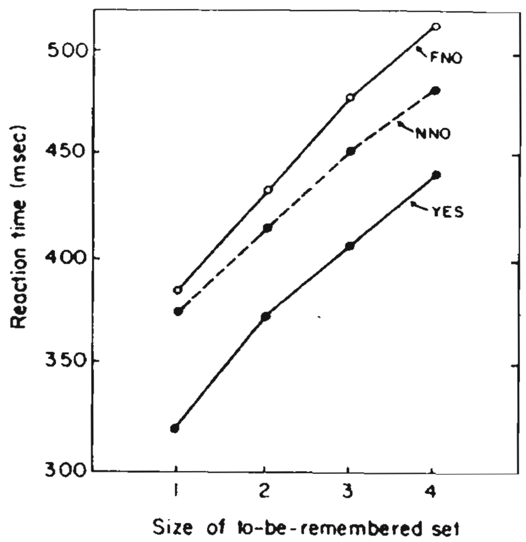

FIG. 16.1. Reaction times as a function of the size of the second (to-be-remembered) set on a given trial. YES and NO probes denote test items that did or did not appear in the TBR set; F-NO and N-NO probes differ in whether they did or did not appear in the first (to-be-forgotten) set. (After Bjork et al., 1970.)

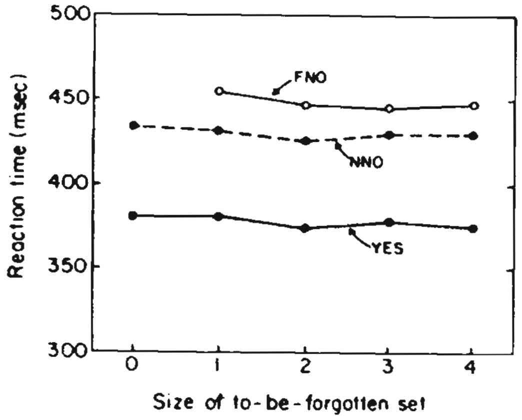

FIG. 16.2. Reaction times as a function of the size of the first (to-be-forgotten) set on a given trial. YES and NO probes denote test items that did or did not appear in the second (to-be-remembered) set; F-NO and N-NO probes differ in whether they did or did not appear in the to-be-forgotten set. (After Bjork et al., 1970.)

Dual-process models of memory scanning, as formulated by Atkinson and Juola (e.g., 1974) and their co-workers (e.g., Juola, Fischler, Wood, & Atkinson, 1971) shortly after the data in Figs. 16.1 and 16.2 were collected, also seemed unable to account for the results. The basic idea underlying such models is that a YES–NO decision can be based either on an initial familiarity judgment or on a systematic search of the type proposed by Sternberg (1966). Test probes whose familiarity is below some low threshold or above some high threshold trigger a rapid NO or YES response, respectively, whereas probes of intermediate familiarity require a slower search process. Response latencies of the former type are presumed to be independent of the size of the search set, whereas response latencies of the latter type are presumed to increase linearly with search set size.

The data in Figs. 16.1 and 16.2 pose two types of problems for such models. First and foremost, given the kind of procedure employed by Bjork et al. (1970) (TBF and TBR sets drawn from the digits 0–9 on every trial), the dual-process model is presumed by Atkinson and Juola (1974) to reduce to a Sternberg-type single-process search model. All items in the set of digits are presented repeatedly across trials, which, according to Atkinson and Juola, renders familiarity judgments essentially useless in distinguishing targets from distractors. Such an assumption enables the dual-process model to account for the greater slopes of response-time functions when items from a small pool of items are reused across trials when compared to the slopes of response-time functions obtained when items from a large pool are used without replacement across trials.

If one attempts to ignore such considerations in an effort to apply the unique-set version of the dual-process model to the present data, a second type of problem emerges. To account for the results in Figs. 16.1 and 16.2, one would like to assume that the prior presentation of F-NO probes in the TBF set increases the familiarity level of such items independent of the TBF and TBR set size on a given trial—resulting in a fixed increment in the likelihood that a search (rather than a rapid NO response) would have to be executed. It could then be assumed that the increased response times for F-NOs (and possibly the greater slope of the response-time function) are a function of the increased likelihood of search for such items. Assuming that such an increment in search likelihood is independent of TBF and TBR set size is, however, completely implausible. As there was no temporal break between the TBF and TBR sets, an F-NO probe on a trial with a small TBR set may be temporally more recent (and, hence, more familiar) than a YES probe on a trial with a larger TBR set. Conversely, with a large TBF set and a large TBR set, a given F-NO probe may have been presented further back in time (in terms of intervening items) than an N-NO probe on a trial when the TBF and TBR sets were small. That is, given that the items used on successive trials were independent samples from the pool of digits, it was frequently the case that an N-NO probe item was presented on the preceding trial as a TBF, TBR, or probe item, and under some conditions that prior presentation may have been quite recent.

There may well be an elaboration of the dual-process model that gives a reasonable account of the data in Figs. 16.1 and 16.2. The point of resurrecting those data, however, is not to argue about alternative scanning models. Rather, the point is the following: We never even entertained an interpretation that is not only entirely consistent with the results, but also consistent with Sternberg's characterization of what slope differences and intercept differences are supposed to reflect. Suppose that TBF items are inhibited by the forget instruction, and that one consequence of that inhibition is to slow down the encoding of those items when they are presented as probes. Once such probes are encoded, suppose further that the subsequent scanning processes are no different from those that follow N-NO probes. Everything in Figs. 16.1 and 16.2 then falls into place: Subjects can limit their search to the TBR set, but F-NO probes take longer to encode, which accounts for the difference in intercept between F-NO and N-NO reaction times.

The present issue is not whether that inhibition-based explanation of the Bjork et al. data is adequate or inadequate when held up to other existing or potential results. In fact, I doubt that such an explanation would hold up in the face of additional experimentation. The point is that we could not, 18 years ago, even generate such an explanation. We exhausted the possible explanations based on computer-type search/scanning/decision processes, but we were unable to even entertain an explanation based on an inhibitory mechanism.

INHIBITION IN DIRECTED FORGETTING

The kind of explanatory bias just illustrated characterized the early work of myself and others on directed forgetting (exceptions being the work of Roediger & Crowder, 1972; Weiner, 1968; and Weiner & Reed, 1969). The effects of a cue to subjects that they can forget the items presented prior to the cue—items that they had been trying, prior to the cue, to learn—are dramatic. Interference owing to the TBF items in the recall of the post-cue TBR items is typically eliminated, and intrusions of TBF items into the recall of TBR items is negligible. It is as if the TBF items had not been presented, although tests of recognition demonstrate that they exist in memory at essentially the same strength as comparable TBR items (although they are nowhere near as recallable). As of 9 or 10 years ago, however, it seemed that all such effects of instructing subjects to forget could be explained in terms of positive actions on the part of the subject—focusing post-cue rehearsal and other mnemonic activities on the TBR items and (somehow) segregating the TBR items in memory in a way that differentiated them from the earlier TBF items.

Those of us working on directed forgetting in that early period (with the exceptions noted previously) were accused by Erdelyi and Goldberg (1979) of having "ignored or brushed aside" (p. 382) the relationships of our work to repression and posthypnotic amnesia. In my own case, I was so convinced that directed-forgetting phenomena had nothing to do with inhibition or repression or posthypnotic amnesia that I would typically start off colloquia with a disclaimer to that effect (sometimes, as a consequence, losing selected members of my audience). Evidence began to accumulate, however, that there was a "missing mechanism" in our explanation of directed-forgetting phenomena (see Bjork, 1978, and Bjork & Geiselman, 1978), and we became convinced, as a consequence of the work reported by Geiselman, Bjork, and Fishman (1983), that the missing mechanism was retrieval inhibition.

Geiselman et al. presented subjects with a list of two different types of items: items subjects were asked to learn and items subjects were asked to judge on a pleasantness scale (incidental items) rather than learn. The two item types were perfectly interleaved in the list (i.e., they were presented in strict alternation). Midway through the list there was an instruction: either that the to-be-learned words presented thus far were practice and should be forgotten, or that the to-be-learned words presented thus far were the first half of the list and should be remembered. After a list of either type, the subjects were asked to free-recall *all* the words they could remember from the list, *judge* words as well as *learn* words, and first-list-half words as well as second-list-half words.

The results of the first of the Geiselman et al. experiments are shown in Fig. 16.3. The learn words show a standard pattern: A cue to forget, compared to a cue to remember, results in lowered recall of the pre-cue (TBF) items (left panel) and enhanced recall of the post-cue (TBR) items (right panel). The important result in Fig. 16.3 is that the incidental (judge) items exhibit the same pattern. For the

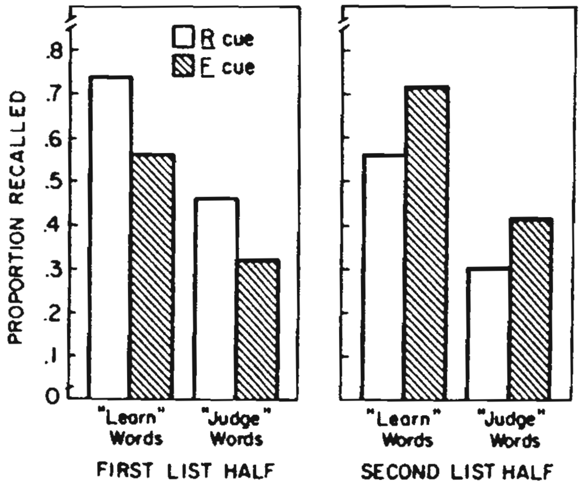

FIG. 16.3. Proportion correct recall as a function of list half and type of midlist cue; R = remember and F = forget. (After Geiselman et al., 1983.)

intentional items, the pattern of results can be explained in terms of differential rehearsal and/or grouping of the second-list-half learn words following the forget cue. The effects of the forget cue on the judge words, however, cannot be explained in those terms. The subjects were not trying to learn or rehearse those items, and, in fact, the forget (or remember) cue was not "aimed" at the judge words. It appears that the judge words, by virtue of their being interlaced with the learn words, are inhibited by the forget cue. They are in the wrong place at the wrong time, so to speak.

The results in Fig. 16.3 can be given two alternative interpretations: (a) that subjects lost track of what items were learn words and what items were judge words, and/or (b) that the pattern in Fig. 16.3 is attributable to differential output interference owing to TBR items being recalled before TBF items. In subsequent experiments, Geiselman et al. were able to rule out those alternatives. They found that subjects were able, with high accuracy, to sort the words they recalled into judge and learn categories. They also found that when the potential for confusion between learn and judge words was eliminated by drawing those words from different categories, the same pattern of results (shown in Fig. 16.3) was obtained. With respect to the possible contributions of differential output interference, Geiselman et al. found that controlling order of output did not change the basic pattern of results.

The inhibition of to-be-forgotten items appears to take the form of retrieval inhibition. When Geiselman et al. tested yes–no recognition rather than recall, none of the variables that make such a difference in Fig. 16.3 mattered at all. The recognition results are shown in Fig. 16.4. The to-be-forgotten learn (or judge) words are as well recognized as the to-be-remembered learn (or judge) words. Although not particularly germane to the present issue, the lack of an overall difference between learn and judge words in recognition is also interesting, given that learn words were better recalled than judge words (see Fig. 16.3). It seems plausible to assume that the two item types may not have differed substantially in depth of processing, but that the learn words—because they were to-be-learned and recalled—received cumulative inter-item associative processing across the list presentation. The effects of any such inter-item processing on recall should be clearly larger than the effects, if any, on recognition.

It is worth commenting that the results obtained by Geiselman et al. do not rule out list differentiation as an important mechanism in directed forgetting. That is, some kind of list or item differentiation is a precondition for TBF items to be selectively suppressed or inhibited. Such an observation is by no means new. Crowder (1976), for example, in his discussion of interference theory, made a similar point in contrasting the response-set suppression (Postman, Stark, & Fraser, 1968) and list-differentiation (Underwood & Ekstrand, 1966, 1967) hypotheses.

It is also important to emphasize, however, that list differentiation alone is not sufficient to explain the present results. Without additional assumptions (such as retrieval inhibition) there is no reason to expect that recall of the first list half in the F-cue condition would be impaired simply by virtue of those items being well differentiated in memory from second-list-half items. In the case of learn words, one can argue that cumulative processing of learn words in the first list half was carried over into the second list half given a midlist R-cue, but that no such carry over occurred given a midlist F-cue. Such "carried over" processing would enhance recall of learn words from the first list half and impair recall of words from the second list half (producing the pattern shown in Fig. 16.3). As stated earlier, however, no such argument can be made to explain the similar pattern observed for the incidental judge words.

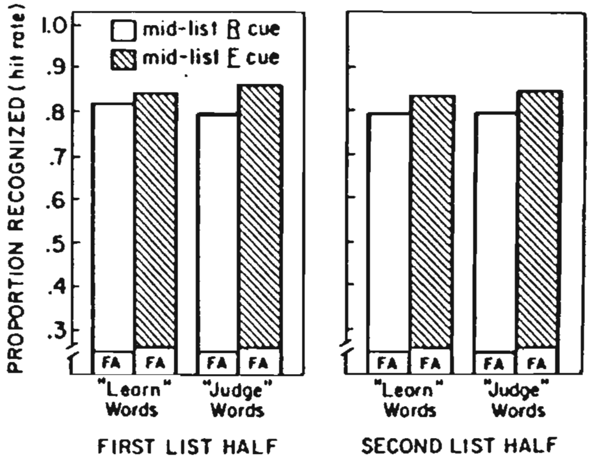

FIG. 16.4. Proportion correct recognition as a function of list half and type of midlist cue; R = remember, F = forget, and FA = false-alarm rate. (After Geiselman et al., 1983.)

RELEASE OF INHIBITION IN DIRECTED FORGETTING

The recognition results shown in Fig. 16.4, and earlier results also showing unimpaired recognition of to-be-forgotten items, indicate that simply presenting the TBF item restores that item to full "strength," as measured by a recognition test. Studies of repetition effects (e.g., Geiselman & Bagheri, 1985), in which TBF items are later represented as TBR items (e.g., in another list), also indicate that TBF items can be readily brought back to full strength—in fact, to approximately the level of items presented twice as TBR items.

Research by Bjork, Bjork, and Glenberg (1973) and Bjork, Bjork, and White (1984) suggests that the absence of interference from TBF items in the recall of TBR items can be changed to interference under certain conditions. Two aspects

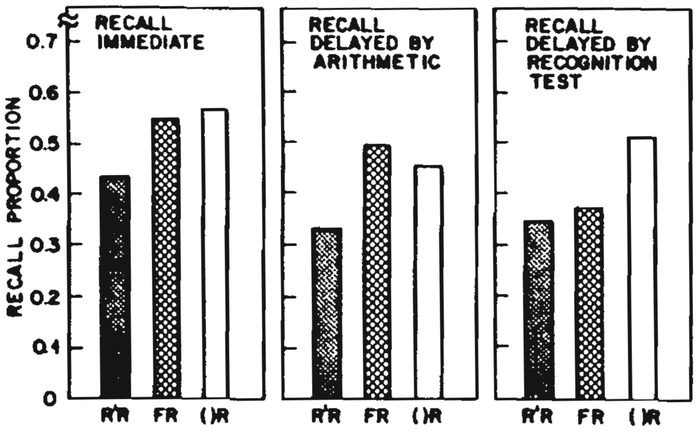

FIG. 16.5. Recall performance on the second sublist of 16 words as a function of list type and test condition; FR and R'R denote lists in which the first sublist of 16 words was to be forgotten and to be remembered, respectively, and (-)R denotes lists where there was no first sublist of words. (After E. Bjork et al., 1973.)

of such reinstatement of interference merit comment. First of all, the simple presence of TBF-item foils on a test of TBR-item recognition causes interference, even though a test of TBR-item recall administered at that same point would not show such interference. Apparently, it is not the case that the lack of interference owing to TBF items in the recall of TBR items is attributable to the subject being able to identify TBF items in memory as TBF items. Rather, it appears that TBF items do not interfere because they are not "encountered" in the recall process—were they to be encountered, they would intrude into and interfere with the recall of the TBR items.

The second way in which the interference owing to TBF items can be reinstated is illustrated by the results shown in Fig. 16.5. E. Bjork et al. (1973) presented subjects with three types of word lists: lists with a midlist cue to forget the first half of the list (FR lists), lists with a midlist cue to remember the first half of the list (R'R lists), and lists without a first half (a shape-judgment task took the place of the words that would have been in the first list half)—(-)R lists. There were 16 words in each list half.

Immediate recall of the words in the second list half is shown in the left panel of Fig. 16.5. Delayed recall of those same words is shown in the middle and right panels, where recall was delayed by an intervening arithmetic task and an intervening recognition test (for second-list-half words), respectively. In the FR and (-)R conditions, subjects were asked to free-recall the words in the second list half; in the R'R condition, they were asked to free-recall all the words in the list. Immediate recall and recall delayed by arithmetic show the same pattern: Recall of second-list-half words is as good in the forget condition as it is in the condition without any first-list-half words, and it is significantly better than recall of those words in the remember condition. When recall is delayed by a recognition test, however, the pattern changes. The forget condition and remember condition are now equivalent, and significantly worse than the condition where there were no first-list-half words.

The recognition test consisted of eight forced-choice tests, four of which paired a TBR word with a TBF word and four of which paired a TBR word with a new word. The recall of the second-list-half items that did not appear on the recognition test is plotted in the right-hand panel. It appears that the exposure of only 4 of the 16 TBF items, in the context of the decision processes involved on the recognition test, was enough to reinstate the interference owing to the entire set of TBF items.

The later results of E. Bjork et al. (1984) add to the results shown in Fig. 16.5 in two important ways. First, they replicate exactly the pattern of results shown in the first and third panels of Fig. 16.5. Second, they demonstrate that it is a necessary condition for the reinstatement of F-item interference that F-items are re-exposed on the recognition test. When recall was delayed via a recognition test that did not include first-list foils, the results were similar to that shown in the middle panel of Fig. 16.5; that is, there was no recovery of the proactive interference owing to F-items.

NECESSARY CONDITIONS FOR RETRIEVAL INHIBITION IN DIRECTED FORGETTING

We have known for some time (see Bjork, 1970) that a cue to forget prior items is much more effective if it precedes the presentation of the to-be-remembered items. If the cue is delayed until after the TBR items are presented (but before they are recalled), the effects of the cue are minimal, no matter how well the TBF and TBR sets are segregated in time and by item type. There is little, if any, reduction in the proactive interference owing to the TBF items on the recall of the TBR items, and the subsequent recall of the TBF items themselves, when requested, is largely unimpaired (in contrast to the near-zero level of recall of those items when the forget cue precedes the presentation of the TBR items).

What we do not know is how to interpret the foregoing pattern of results. It could mean that TBF items are inhibited as a consequence of learning the subsequent TBR items (that is, of starting the learning process over again). It could also mean, however, that retrieval inhibition is a direct consequence of being instructed to forget, but that such instructions can be delivered too late to be effective (owing, perhaps, to consolidation-type processes that would stabilize the representation of TBF items in memory).

Gelfand and Bjork (1985) set out to address that issue by means of the

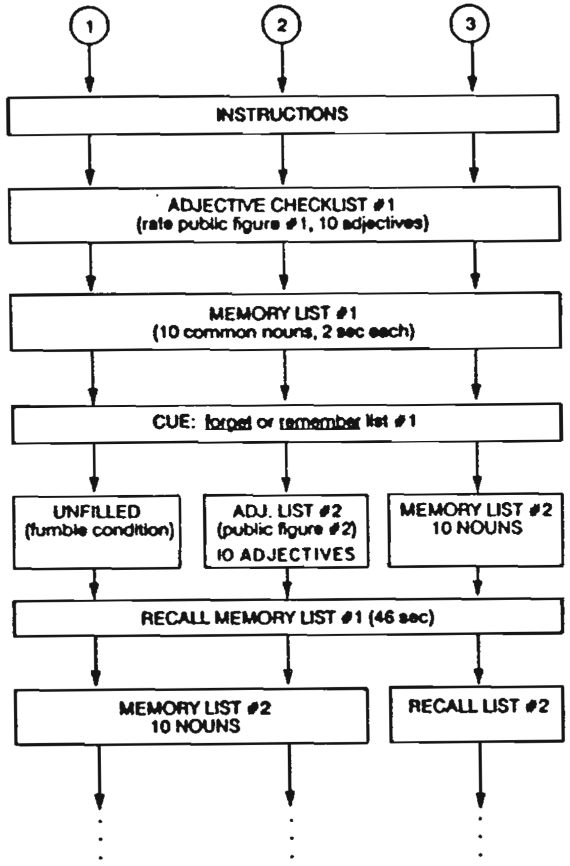

FIG. 16.6. Sequence of events for groups of subjects that differed in whether they were cued to forget or remember a first list of nouns and in the nature of the activity interpolated between that cue and a test of their ability to recall those nouns. (After Gelfand & Bjork, 1985.)

experiment diagrammed in Fig. 16.6. The subjects knew that they were going to be presented two types of 10-word lists: lists of 10 adjectives, about which they were to say whether each adjective in turn was or was not descriptive of a certain public figure (e.g., Ronald Reagan), and lists of 10 nouns, which they were to try to remember. After going through a first adjective check list, and being presented a first list of nouns to learn, half of the subjects were then cued to forget the first list of nouns ("what you have done so far is practice . . .") and the remaining subjects were cued to remember the first list of nouns. The critical manipulation, as shown in Fig. 16.6, is what happened next. One group of subjects did nothing (while the experimenter fumbled around killing time), another group was presented a second adjective check list, and the third group learned a second list of nouns. After that intervening activity, all groups were tested for their recall of the first list of nouns. Groups 1 and 2 were then presented a second list of nouns to learn, whereas group 3, which had already been presented a second list of nouns, was asked to recall that second list. The subsequent events in the design are not relevant to the present issues.

TABLE 16.1
Proportion List-1 Words Recalled as a Function of the Post-List-1 Cue and the Activity Interpolated Prior to the Recall of List 1 (after Gelfand & Bjork, 1985)

	Interpolated Activity		
Cue re memory list 1	*Unfilled*	*Adjective Checklist*	*Memory List 2*
Forget	.58	.53	.37
Remember	.61	.55	.54

The proportions of list-1 words recalled in the various conditions are shown in Table 16.1. The pattern of results in Table 16.1 seems to argue strongly that the retrieval of TBF items is inhibited as a by-product of new learning. Compared to the remember-cue conditions, there was only minor loss of list-1 recall owing to the forget cue in the unfilled and adjective-check-list conditions (.58 vs. .61 and .53 vs. .55, respectively). When the second to-be-memorized list was interpolated, however, there was substantial loss of access to list 1 (.37 vs. .54). Were the forget cue to operate as a kind of magic wand, inhibiting the TBF items as a direct consequence of the instruction itself, there should be no interaction of the type shown in Table 16.1. That is, it is the resetting of the learning process initiated by the presentation of a new list of to-be-remembered items that is necessary for the prior to-be-forgotten items to be inhibited.

An additional result, consistent with the results of E. Bjork et al. (1973, 1984), deserves mention. In the F-cue condition, the effort by subjects to recall the first list reinstated the interference owing to those items in the recall of the second list. In the third column of Fig. 16.6, the levels of recall of list 2 (following the effort to recall list 1) were .48 and .46 in the F-cue and R-cue conditions, respectively.

RETRIEVAL INHIBITION IN THE BROADER CONTEXT

Updating

The layperson tends to think that the singular problem in using one's memory effectively is to remember things. Efficient use of our memories, however, also requires that we update the system effectively. When we need to remember our current phone number, the current married name of a female acquaintance, the trump suit on this hand, and where we left the car today, it does not help us to remember our old phone numbers, our friend's maiden name, the trump suit on the last hand, and where we left the car yesterday. That is, we need some means to suppress, set aside, destroy, or discriminate out-of-date information in memory in order to remember current information effectively.

The updating process in computers is brutal but effective: When a new value of some variable is entered, the old value is destroyed. There are clear drawbacks to such a system, however, should the old value become pertinent once again in some way. Retrieval inhibition, as an updating mechanism, is a more sophisticated system with the following adaptive properties: (a) Because the old material becomes nonretrievable (by virtue of learning the new material), that material also becomes non-interfering in the recall of the new material; (b) the old material, however, remains in memory, is apparently at full strength from a recognition standpoint, and is, therefore, familiar and identifiable when it reoccurs; and (c) the old material is not only recognizable but also, apparently, relearnable in the sense that it becomes fully accessible in memory when presented again as to-be-learned material.[1]

Unitization

A recent dissertation by Hirshman (1988) suggests that retrieval inhibition may play an adaptive role in another context. In a large number of experiments, Hirshman explored a paradoxical effect of semantic relatedness. When subjects were asked to study a list of word pairs differing in associative strength (some strongly related, like TABLE–CHAIR, some weakly related, like GLUE–CHAIR), their ability to free recall the response words was actually better for the weakly related pairs. Such a result is paradoxical on several grounds. First, cued recall, not surprisingly, shows a large advantage for strongly related pairs. Second, free recall of the response component of a pair is often going to be mediated by recall of the stimulus component, and such mediation should clearly be more effective the stronger the associative relationship between the stimulus and the response terms. Finally, the advantage of weakly related pairs in free recall goes against a strong generalization in verbal learning: that learning is positively related to meaningfulness. The kind of results obtained by Hirshman are shown in Table 16.2.

---

[1] Donald Broadbent (personal communication) has pointed out that, in some respects, computers are not as different from human memory as this passage suggests. In typical file management systems, an instruction to erase a file does not destroy the information in that file, but, rather, changes entries in a directory; the locations where that file was stored become available for later storage. Until those locations are overwritten, the "erased" file can be retrieved by special means—but *only* by special means, which Broadbent argues may be analogous to a kind of retrieval inhibition. Even when an updated (revised) version of a given file is stored, the old version may or may not be destroyed, depending on the design of the system. What can be said with some confidence, Broadbent's interesting arguments notwithstanding, is that the write process in a computer is fundamentally destructive in a way that the write process in humans is probably not.

TABLE 16.2
Proportion of Response Words Recalled as a Function of Type of Test and Associative Strength (after Hirshman, 1988, Experiment 1)

Type of Test	*Strongly Related Pairs*	*Weakly Related Pairs*
Free recall	.23	.34
Cued recall	.93	.73

Based on substantial empirical support, Hirshman isolated encoding and retrieval processes that he believes give rise to such paradoxical effects of semantic relatedness. My own less-well-supported view of his research is that such effects illustrate a fundamental associative principle that characterizes human memory (and the development of expertise): as items become associated to the point that they become unitized, independent access to those items becomes inhibited. The ability to operate in higher-order units composed of components that were once separate units (in children, e.g., going from letters as units to words as units) is critical to the development of intelligence and skills. Apparently, inhibiting access to the separate components of a higher-order unit is adaptive in the sense that it is part and parcel of those components acting as one chunk in memory.

A colleague in my department at UCLA (Wendell Jeffrey) mentioned an anecdote that is probably a good everyday example of the unitization effect. After ordering an item from a catalog by phone, he was asked for his name and address by the operator. As he was starting to give what for him, and for most of us, is a fairly unitized sequence, the operator said she needed his zip code first. He momentarily was unable to retrieve that overlearned item and had to go through the sequence to get it. When he apologized to the operator, she said "that happens all the time; I don't know why the computer wants it that way."

Memory for Serial Order

The difficulty that Wendell Jeffrey experienced in attempting to go directly to the last component of his address, also illustrates the possible role of inhibition in the retention of order information—as postulated by Estes (1972). Estes assumes that for the items in a sequence to be output properly ". . . the individual must inhibit response tendencies to later items in the sequence until the responses to earlier items have been emitted" (p. 183). As a consequence of rehearsal and practice, in his theory, "the inhibitory tendencies required to properly shape the response output become established in memory and account for the long-term preservation of order information" (p. 183). More specifically, he assumes that when the higher-order "control element" that stands for the sequence in memory is activated, excitation flows to all elements in the sequence. However, the responses associated with later elements in the sequence are inhibited, owing to the inhibitory connections that have been established from each earlier item to each later item. Upon the actual output of an earlier item, Estes assumes, the inhibitory input from the item to later items ceases.

Another role that must be played by some kind of inhibitory mechanism in the output of an ordered sequence is the inhibition of items in the sequence that have already been generated. That is a particularly vexing problem conceptually, because—as is discussed in the next section—generating an item makes that item more accessible in memory and may block access to related items in memory.

TYPES OF RETRIEVAL INHIBITION

As mentioned at the outset, the word *inhibition* is used in a number of ways. I have focused on inhibition as a suppression-type process directed at the to-be-inhibited information for some adaptive purpose. I assume that the target of such suppression is information that already resides in memory, hence, the term ***retrieval*** inhibition. Such a usage is close to inhibition in the classic sense of ***repression*** (cf. Freud, 1914), or the class of processes Osgood (1953) refers to as "motivated forgetting."[2] I refer to the "classic sense" of repression because, in more recent times, repression is often defined as an unconscious defense mechanism, whereas retrieval inhibition of the type referred to here would, typically, be characterized by a conscious intent to suppress. In their very thorough discussion of the meaning of repression, however, Erdelyi and Goldberg (1979) point out that Freud himself viewed repression as typically intentional and "from his earliest psychological writings to his last . . . uses 'suppression' and 'repression' interchangeably" (p. 366).

It is retrieval inhibition in the foregoing sense that has played little role in our recent theories of memory. Another type of "retrieval inhibition"—when retrieval of target information from memory is blocked temporarily by the retrieval of related non-target information—has played a much more frequent role in our theorizing (see, e.g., Bjork & Geiselman, 1978; Blaxton & Neely, 1983; A. Brown, 1981; J. Brown, 1968; Kato, 1985; Nickerson, 1984; Roediger, 1974, 1978; Roediger & Neely, 1982; Rundus, 1973; Tulving & Hastie, 1972; and Watkins, 1975). Such inhibition or blocking comes about, presumably, because there is a kind of capacity limitation on retrieval: As some items in a category are activated in the sense of being made more accessible in memory—via their being presented or recalled—the access to other items is slowed down or blocked. There may be good reasons that such retrieval blocking is a more frequent construct in our theories, but its more frequent use may be attributable, in part, to its being more compatible with the computer metaphor. Retrieval inhibition/suppression and retrieval inhibition/blocking share the basic properties that access to some target information in memory is impaired, and that such impairment is temporary, but differ in most other important respects.

Retrieval Inhibition/Suppression

Inhibition in the sense of suppression would seem to have the following characteristics: It is directed at the to-be-inhibited information, and it is initiated to achieve some goal (such as the reduction of proactive interference, or the avoidance of painful recollection). In addition to some of the examples discussed previously, the ***response-set suppression*** hypothesis of Postman et al. (1968) is a perfect example of an inhibitory process with those properties. According to Postman et al., when a second to-be-learned list is similar to a first (e.g., in the A-B, A-D paradigm), retroactive interference is, in part, attributable to subjects suppressing first-list responses during second-list learning. Such suppression supposedly facilitates second-list learning by blocking the covert or overt intrusions of first-list responses during second-list learning, but it also leads to worsened performance on list 1, should that list be tested after list-2 learning is completed. Such suppression is directed at the to-be-suppressed items, and it is adaptive in the sense that it facilitates list-2 learning.

Retrieval Inhibition/Blocking

Inhibition in the sense of blocking, on the other hand, is a by-product of the activation of other items in memory, and may not be adaptive—at least on the short term. Consider the kind of retrieval blocking demonstrated by J. Brown (1968): Subjects who spent 5 minutes studying the names of 25 of the 50 states later recalled more of those 25 states but fewer of the other 25 states than did subjects in a control group that spent those 5 minutes doing light reading. It was not the goal of the experimental subjects to inhibit the 25 names of states they did not study, nor was such inhibition adaptive in terms of the goal of recalling all 50 states, nor were the pre-recall activities of those subjects directed at those 25

---

[2] With respect to my earlier point that the computer metaphor is partly responsible for our not appealing to inhibitory processes as explanatory constructs, it is of historical interest to note that the handbooks of experimental psychology written by Osgood (1953) and by Woodworth and Scholsberg (1954) cover "motivated forgetting" in some detail. Those thorough and highly rigorous books were written prior to the onset of the information-processing approach in the late 1950s and its accompanying computer metaphor. The stimulus–response approach was still dominant in the early 1950s, and notions of inhibition, suppression, unlearning, and so forth were more compatible with that approach.

names, but, rather, at the 25 names they *did* study. (For a review of the variety of such part-list cuing procedures, and the negative effects of such procedures on the recall of uncued items, see Nickerson, 1984).

Complications and Gray Areas

In principle, the properties of inhibition as suppression and inhibition as blocking seem fairly well defined and non-overlapping. In practice, however, it is easy to cite phenomena and theories that illustrate that those two types of retrieval inhibition are neither exhaustive nor mutually exclusive. Consider, for example, the Gelfand and Bjork (1985) results shown in Table 16.1. As discussed earlier, those results seem to be strong evidence that the to-be-forgotten list is inhibited as a by-product of learning the new list—not simply as a consequence of the forget instruction itself. That sounds like inhibition/blocking is the operative mechanism, but then one might have expected the middle condition in Fig. 16.6, in which ten adjectives were "studied" prior to list-1 recall, to have resulted in some inhibition of the to-be-forgotten words as well. It could be that both the intent to update the system (that is, to forget some prior information) and the process of storing/activating new information are necessary conditions for retrieval inhibition. Without the ongoing list-2 learning, for example, the suppression of list-1 responses (as in the response-set suppression hypothesis of Postman et al., 1968) may not be possible. Even repression, in the psychoanalytic sense, might be accomplished by people activating more pleasant memories rather than directly suppressing the offending memory.

The *unlearning hypothesis* put forth by Melton and Irwin (1940) as a factor in retroactive interference seems like a similar ambiguous case. Is one "unlearning" a specific list-1 response to a given stimulus during list-2 learning by suppressing that response or by storing a new response that blocks access to the list-1 response? Once again, both the intent to suppress and the opportunity to store a new association may be necessary.

The possible role of retrieval inhibition in unitization and the retention of serial order illustrates another problem. The inhibition of direct access to individual elements in a well-learned sequence may be adaptive, as argued previously, but it is difficult to characterize the process leading to such inhibition as intentional on the part of the subject. Such inhibition, which seems more like suppression than blocking, also seems more like an automatic by-product of other associative activities in the brain, rather than a process over which the organism has control.

What is adaptive and is not adaptive, on the short term and on the long term, is also not so easy to judge. In the J. Brown (1968) example cited earlier, the blocking of later retrieval of the non-studied states by the studied states did not seem adaptive, and was not intentional on the subjects' part. Overall, however, retrieval blocking may be an adaptive component in updating the system—that is, in facilitating access to current information. Information that is accessed in memory becomes more accessible (a new home phone number, e.g.), and information that is not accessed, even if perfectly accessible at some earlier point in time (an old home phone number, e.g.), gradually becomes less accessible. In general, such a system will result in information that is more pertinent to the present state of affairs being more accessible in memory owing to its having been retrieved more often in the recent past. Information that is less pertinent, or completely out of date, will be non-retrievable and, hence, non-interfering. (For a more detailed argument regarding the adaptive side of retrieval failure in autobiographical memory, see Bjork & Bjork, 1988).

Finally, there are a variety of recent phenomena that can be interpreted either in terms of retrieval/blocking mechanisms or in terms of retrieval/suppression mechanisms. Blaxton and Neely (1983), for example, in discussing inhibition in the cued recall of category exemplars owing to the prior generation of other exemplars (measured in terms of reaction times), outline two theoretical alternatives, a retrieval/blocking interpretation derived from the Raaijmakers and Shiffrin (1981) model (SAM), and a retrieval/suppression interpretation derived from the Keele and Neill (1978) model of attention.

"Retrieval" Inhibition at Encoding

This chapter has focused on inhibition in the retrieval of information from memory. I have not addressed the formidable array of evidence for inhibition in sensory and perceptual processes, in attention, in lexical decision tasks, and in recognition tasks. On the surface, at least, such inhibition would seem to impair an encoding and/or judgment process—not a retrieval process. The encoding/retrieval distinction is not an entirely clean one, however, as my earlier discussion of the Bjork, Abramowitz, and Krantz (1970) results indicates (Figs. 16.1 and 16.2). The inhibition-based interpretation of those results assumes that the encoding of a negative probe is slowed down (inhibited) when that probe has appeared earlier in the to-be-forgotten set. In that interpretation, the slowed encoding of the probe is attributed to the retrieval inhibition effected by the forget cue. Similarly, Neely and Durgunoğlu (1985) account for inhibited (slowed) episodic recognition of words following a semantically related prime in terms of a suppression-type retrieval inhibition. They assume that "subjects actively tried to suppress the prime's semantic associates in the episodic recognition task because the retrieval of such associates was irrelevant to the task—namely, knowing that a target is semantically related to its prime does not provide information as to its study status" (p. 485).

These two examples, as well as others that could be cited (e.g., Dosher, 1984), illustrate that retrieval inhibition may impair a subject's ability to encode or recognize or judge the lexical status of an inhibited item. Those examples also suggest that inhibition on the input side may be more active/volitional than is often thought to be the case. An additional striking example that underscores that point comes from research by Allport, Tipper, and Chmiel (1985). They presented subjects with trials consisting of a prime stimulus followed by a probe stimulus. Both stimuli consisted of a red letter and a green letter on top of each other (the prime might be a red A and a green B, for example, and the probe might be a red C and a green D). At the end of a trial, subjects were asked first to name the red letter in the probe stimulus and then to recall the red letter from the prime stimulus. On the trials where the green (to-be-ignored) letter in the prime stimulus was repeated as the red (to-be-named) stimulus in the probe stimulus, there was an inhibitory effect: The subjects were slowed down in their ability to name that letter. On the final trial of the experiment, subjects—rather than being presented a probe stimulus—were asked to recall the green (to-be-ignored) letter from the preceding prime stimulus. The fact that few subjects could do so indicates that a kind of active retrieval inhibition was responsible for the slowed encoding observed by Allport et al. (1985).

CONCLUDING COMMENT

In all likelihood, the general thesis in this chapter—that inhibitory processes play a critical role in the overall functioning of human memory—will seem uncontestable in the near future, possibly even by the time this chapter appears in print. A new metaphor has emerged to influence the thinking of memory researchers: the brain metaphor. As other chapters in this volume attest, our ideas about memory are being shaped by the accelerating knowledge of the possible functions of certain brain structures in human memory. In addition, the neural/connectionist approach to the simulation of cognitive processes is being pursued with enthusiasm and high expectations. In contrast to the computer metaphor, which tended to lead us away from explanations based on inhibitory processes, the brain metaphor, if anything, will push us towards such explanations.

ACKNOWLEDGMENTS

Preparation of this essay was aided by a grant from the Academic Senate of the University of California. The author thanks James Neely, Larry Jacoby, Elizabeth Bjork, and Tom Minor for useful comments and criticisms, and Henry Roediger for his unsparing efforts to edit this chapter.

REFERENCES

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

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (pp. 243–293). San Francisco: Freeman.

Bjork, E. L., & Bjork, R. A. (1988). On the adaptive aspects of retrieval failure in autobiographical memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1: Memory in everyday life* (pp. 283–288). London: Wiley.

Bjork, E. L., Bjork, R. A., & Glenberg, A. (1973, November). *Reinstatement of interference owing to to-be-forgotten items*. Paper presented at the meeting of the Psychonomic Society, St. Louis, MO.

Bjork, E. L., Bjork, R.A., & White, S. (1984, November). *On the induced recovery of proactive interference*. Paper presented at the meeting of the Psychonomic Society, San Antonio, TX.

Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior, 9*, 255–268.

Bjork, R. A. (1978). The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, (Vol. 12, pp. 235–259). New York: Academic Press.

Bjork, R. A., Abramowitz, R. L., & Krantz, D. H. (1970, April). *Selective high-speed scanning of item-sets in memory*. Paper presented at the Midwest Mathematical Psychology Meetings, Bloomington, IN.

Bjork, R. A., & Geiselman, R. E. (1978). Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 344–361.

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. (1981). Inhibition in cued retrieval. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 204–215.

Brown, J. (1968). Reciprocal facilitation and impairment of free recall. *Psychonomic Science, 10*, 41–42.

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

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed–accuracy study. *Cognitive Psychology, 16*, 519–555.

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 Associates.

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*. Washington, DC: Winston.

Freud, S. (1914). *Psychopathology of everyday life*. New York: Macmillan.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory & Cognition, 13*, 57–62.

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.

Gelfand, H., & Bjork, R. A. (1985, November). *On the locus of retrieval inhibition in directed forgetting*. Paper presented at the meeting of the Psychonomic Society, Boston, MA.

Hirshman, E. L. (1988). The expectation–violation effect: Paradoxical effects of semantic relatedness. *Journal of Memory and Language, 27*, 40–58.

Juola, J. F., Fischler, I., Wood, C. T., & Atkinson, R. C. (1971). Recognition time for information stored in long term memory. *Perception & Psychophysics, 10*, 8–14.

Kato, T. (1985). Semantic-memory sources of episodic retrieval failure. *Memory & Cognition, 13*, 442–452.

Keele, S. W., & Neill, W. T. (1978). Mechanisms of attention. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception*, (Vol. 9, pp. 3–47). New York: Academic Press.

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, 53*, 173–203.

Neely, J. H., & Durgunoğlu, A. Y. (1985). Dissociative episodic and semantic priming effects in episodic recognition and lexical decision tasks. *Journal of Memory and Language, 24*, 466–489.

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

Osgood, C. E. (1953). *Method and theory in experimental psychology*. New York: Oxford.

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

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

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition, 2*, 261–296.

Roediger, H. L. (1978). Recall as a self-limiting process. *Memory & Cognition, 6*, 54–63.

Roediger, H. L., & Crowder, R. G. (1972). Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology, 3*, 255–267.

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

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

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652–654.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421–457.

Tulving, E., & Hastie, R. (1972). Inhibition effects of intralist repetition in free recall. *Journal of Experimental Psychology, 92*, 297–304.

Underwood, B. J., & Ekstrand, B. R. (1966). An analysis of some shortcomings of the interference theory of forgetting. *Psychological Review, 73*, 540–549.

Underwood, B. J., & Ekstrand, B. R. (1967). Studies of distributed practice, XXIV: Differentiation and proactive inhibition. *Journal of Experimental Psychology, 74*, 574–580.

Watkins, M. J. (1975). Inhibition in recall with extralist "cues." *Journal of Verbal Learning and Verbal Behavior, 14*, 294–303.

Weiner, B. (1968). Motivated forgetting and the study of repression. *Journal of Personality, 36*, 213–234.

Weiner, B., & Reed, H. (1969). Effects on the instructional sets to remember and to forget on short-term retention: Studies of rehearsal control and retrieval inhibition (repression). *Journal of Experimental Psychology, 79*, 226–232.

Woodworth, R. S., & Scholsberg, H. (1954). *Experimental psychology*. New York: Holt.

17

Theoretical Issues in State Dependent Memory

Eric Eich
University of British Columbia

The idea that what has been learned in a certain state of mind or brain is best remembered in that state is an old and familiar one in psychology. The credit for this concept—one that I refer to as *state dependent memory*—goes to an astute French aristocrat, the Marquis de Puységur (Chastenet de Puységur, 1809; Ellenberger, 1970). In 1784, Puységur discovered that although a person might appear, upon awakening, to be amnesic for events that had occurred during hypnosis, memory for these events returned once the individual reentered a state of "magnetic sleep"—Puységur's term for hypnosis. Decades later, a French physician named Azam (1876) related a strikingly similar observation in connection with the case of a young woman who suffered sudden attacks of hysterical somnambulism, or "pathological sleep," as the disorder was then known. And in an article published in 1910, Morton Prince conjectured that the reason most people have difficulty remembering their dreams is not because they do not *want* to remember—as Freud (1953) and other psychodynamicists of the day were claiming—but rather, because they *cannot* remember, owing to the dissimilarity between the states of "natural sleep" and ordinary wakefulness.

Like psychology itself, the concept of state dependent memory has a long past, but a short history as a subject of scientific study. Indeed, the earliest experimental demonstrations of human state dependence date only to the late 1960s (see Eich, 1977; Weingartner, 1978), and of the 100-plus empirically oriented articles of the subject that have appeared since then, approximately 70% have been published within the last 10 years.

During the first decade of research on human state dependent memory, from about 1967 to 1976, almost all experiments dealt with drug-defined differences in state. This type of experiment is exemplified by the work of Goodwin, Powell,