Journal of Experimental Psychology: General
1983, Vol. 112, No. 1, 58–72

0096-3445/83/1201-0058$00.75

Disrupted Retrieval in Directed Forgetting: A Link With Posthypnotic Amnesia

Ralph E. Geiselman, Robert A. Bjork, and Deborah L. Fishman
University of California, Los Angeles

Certain reliable findings from research on directed forgetting seem difficult to accommodate in terms of the theoretical processes, such as selective rehearsal or storage differentiation, that have been put forward to account for directed-forgetting phenomena. Some kind of "missing mechanism" appears to be involved. In order to circumvent the methodological constraints that have limited the conclusions investigators could draw from past experiments, a new paradigm is introduced herein that includes a mixture of intentional and incidental learning. With this paradigm, a midlist instruction to forget the first half of a list was found to reduce later recall of the items learned incidentally as well as those learned intentionally. This result suggests that a cue to forget can lead to a disruption of retrieval processes as well as to the alteration of encoding processes postulated in prior theories. The results also provide a link between intentional forgetting and the literature on posthypnotic amnesia, in which disrupted retrieval has been implicated. With each of these procedures, the information that can be remembered is typically recalled out of order and often with limited recollection for when the information had been presented. It therefore was concluded here that retrieval inhibition plays a significant role in nonhypnotic as well as in hypnotic instances of directed forgetting. The usefulness of retrieval inhibition as a mechanism for memory updating was also discussed.

Although most instances of forgetting are unintentional or incidental, there are occasions when we try to forget, either because the memory is unappealing or because the memory constitutes a source of interference in conducting routine mental operations such as memory updating. Consequently, selective forgetting has been assigned an integral role in the processing of to-be-remembered (R) information by various scholars and researchers. In 1882, for example, Ribot wrote in his book that

> without the total obliteration of an immense number of states of consciousness, and the momentary repression of more, recollection would be impossible. Forgetfulness, except in certain cases, is not a disease of memory, but a condition of its health and life. (p. 61)

In 1890, James wrote that "if we remembered everything, we should on most occasions be as ill off as if we remembered nothing" (p. 680). More recently, Bjork (1972) has stated, "That we need to update our memories is clear: We would degenerate to a proactive-interference-induced state of total confusion otherwise" (p. 218).

The apparent importance of selective forgetting in daily experience has led to a broad-based search for the mental mechanisms underlying intentional forgetting. Some memory theorists have contended that intentional forgetting can be carried out simply through selective inattention. The *Roth Memory Course* (Roth, 1918/1961), for example, teaches that

> by denying any attention to your temporary mental [associations] after they have served their purpose, they will pass out of your mind. In this way, your [coding schemes] are left free for filing other facts. (p. 287)

There is some empirical evidence, however, that subjects can influence the inaccessibility of memories to a greater extent with a deliberate motivational set to forget than with a passive nonrehearsal or inattention strategy (Weiner & Reed, 1969). This article attempts to (a) evaluate the explanatory status of certain mechanisms that have been proposed to

Portions of this research were presented by the senior author as part of a symposium on directed forgetting at the meeting of the American Psychological Association, Toronto, August 1978.

Requests for reprints should be sent to Ralph E. Geiselman, Department of Psychology, University of California, Los Angeles, California 90024.

account for directed-forgetting phenomena, (b) suggest an alternative methodology to circumvent some of the methodological limitations that characterize past research, (c) provide evidence that retrieval inhibition is an important mechanism in the intentional forgetting of information, and (d) note an empirical link between two currently independent literatures on directed forgetting.

Directed Forgetting: Basic Findings

One popular method for studying the memory mechanisms involved in selective forgetting has been to instruct subjects at some point during the presentation of R items that the items already presented are now to be forgotten and will not be tested later. The volume of literature that has been generated under the rubric of directed forgetting has been enormous with an incredible number of variations in procedure (see Bjork, 1972; Epstein, 1972; and Timmins, 1976, for reviews). In many of the more recent experiments, these paradigms have been used as investigative tools to address issues of memory other than the nature of intentional forgetting itself. Among the topics that have been addressed are effects of rehearsal on recall and recognition (Woodward, Bjork, & Jongeward, 1973) and constituent processes in the differentiation of items in memory (Bjork & Geiselman, 1978).

For the present purposes, there are two things that we would like to know in considering studies of directed forgetting. The first is the effect of the forget instruction on the recall of the R items. Hence, the typical study compares the level of recall of R items presented after a cue to forget the prior items in a list with the level of recall of R items in two control conditions—one in which only the R items are presented, and one in which subjects are not given the cue to forget the initial items. Such comparisons typically demonstrate the striking power of the forget instruction to eliminate the interference from the to-be-forgotten (F) items—sometimes completely (Bjork, 1970). Studies of directed forgetting, then, substantiate the suggestions of Ribot (1882), James (1890), and Bjork (1972) that selective forgetting plays an integral role in the successful processing of R information.

The second thing one would like to know is the fate of those items that have been cued to be forgotten. This question is the primary focus of the present research. Even though the recall of F items is typically very low when tested, the recognition of F items is typically quite high (Davis & Okada, 1971), sometimes equal to that for R items (Block, 1971; Geiselman, 1974, 1977). Because of these results, the hypothesis that a cue to forget serves to erase an F item from memory has been dismissed. Similarly, it does not appear that a cue to forget inhibits the consolidation of an F item in memory (Bjork & Geiselman, 1978).

Theories of directed forgetting largely have tended to emphasize selective remembering, not selective forgetting. That is, the difference in a subject's ability to recall the R versus F items has usually been explained in terms of a difference in the amount or type of rehearsal given to the items. There are a variety of results that are consistent with such an explanation. Often, for example, the increase in recall of R items that accompanies a cue to forget is approximately equal in magnitude to the corresponding decrease in recall of the F items (Geiselman, 1974; Reitman, Malin, Bjork, & Higman, 1973). Also, if the forget instruction is given at the time of recall rather than immediately following the F items, the forget instruction has little if any effect on recall performance (Bjork, 1970). Finally, one manipulation that appears to reduce the deleterious effect of cue to forget on F-item recall is to increase the amount of elaborative rehearsal given to the item before it is cued to be forgotten. For example, Bugelski (1970) has shown that requiring complex processing of F words, such as imaging, eliminates the effect of forget cues on the recall of the F words. Also, Roediger and Crowder (1972) have shown that when subjects are told to forget an item within the Brown–Peterson paradigm, the subjects perform more optimally on the distractor task.

Bjork (1970, 1972) has suggested that two interrelated processes operating at encoding can account for most of the observed directed-forgetting phenomena. According to Bjork's theory, subjects are presumed to (a) devote all post-F-cue rehearsal and other mnemonic activities to the R items but also to (b) group R items in memory in a way that functionally segregates or differentiates them from the F items. The addition of the differential-grouping mechanism is necessary to account for two classes of results. One is that when paired associates are used as the stimulus materials, subjects rarely give R responses to F-item probes or vice versa, even though they do make intrasublist intrusion errors (Bjork, 1970). The other is that if a series of logically connected sentences is used as the stimulus materials, the decrease in F-item recall and the increase in R-item recall are less than if sentences are used that do not have a logical ordering (Geiselman, 1974). This result implies that differential grouping of the items is necessary in order to conduct selective rehearsal efficiently.

A Missing Mechanism?

Although the selective-rehearsal and differential-grouping mechanisms together give a reasonable qualitative account of directed-forgetting phenomena, there are some situations where it is difficult to explain how subjects could accomplish either selective rehearsal or differential grouping. Consider one example. In an experiment by Jongeward, Woodward, and Bjork (1975), subjects were presented lists of 32 words, each of which consisted of eight four-word sets. Each set of four words was presented one word at a time (2.3 sec per word). After the last word in a set, there was a 3-sec rehearsal period, which was in turn followed by a 1-sec cue that informed the subjects as to which of the four words, if any, they would have to remember at the end of the experiment. Immediately after the cue, the next word set was presented. Subjects were instructed to devote all of their rehearsal activities to the current set of words, something they said they were only too happy to do because of the demanding nature of the task.

Jongeward et al.'s experimental procedure would appear to provide little in the way of opportunities to devote differential processing to R versus F words. During the 12.2 sec from the time the first word in a set was presented to the cue following the rehearsal period, subjects did not know which, if any, of the four words they would need to remember. Subjects were asked not to rehearse words from prior sets during the presentation of later sets, and they said that they did not. In support of their claim, there was no relation obtained between the probability of recalling an R word and the number of R words in the following set. In spite of all that, subjects recalled approximately 35% of the R words but intruded less than 5% of the F words.

Another indication that subjects may use an F cue to do more than simply stop processing the F items can be seen in a study of pupillary responding during a directed-forgetting task. The magnitude of task-evoked pupillary dilations is commonly taken as an index of cognitive load. Johnson (1971) reported a short-lived increase (followed by a sizeable decrease) in pupil dilation in subjects immediately following a cue to forget all prior items in a list. Although it has been argued that the momentary decrease in residual processing capacity is due to the processing of the cue itself, this seems unlikely. It should not take more processing capacity to process a cue that is well learned in the experiment and expected occasionally than to process another list item. Further, it is probably not the case that the pupillary response immediately following an F cue is due solely to the preparation of a new rehearsal scheme for the R items because such a process is also necessary at the very beginning of the list where no such response was observed. A startle-response explanation is inappropriate as well because a control group who received the cues without meaning did not show a corresponding pupil dilation. It is curious then what must have occurred, either consciously or unconsciously, immediately following the cue to forget.

Retrieval Inhibition

The results mentioned above and some results from additional experimental paradigms (cf. Bjork & Geiselman, 1978) suggest that there may be a missing mechanism in the account of directed forgetting. One possibility is that a cue to forget can initiate a process that inhibits or blocks access routes to the episodic memory traces corresponding to the F items, making them nonretrievable at the time of recall except in the presence of prepotent retrieval cues, such as the copy cues on a recognition test. With respect to the retention of items that are explicitly cued to be forgotten, we have not had an adequate paradigm to evaluate the notion of retrieval inhibition. Within the framework of directed forgetting, if one were to instruct a subject that he should actively attempt to forget (repress) the F items, how then should we measure his retention of those items? We have created a dilemma for our subjects. Our results may be due to nothing more than conscious response withholding (suppression) of retrievable items. As Weiner (1968) has noted,

> one is reminded (in such situations) of a story about a king who was told that any wish would be granted, as long as he did not think of the eye of a camel when making the wish. Needless to say, none of his wishes were fulfilled. (p. 217)

Similarly, Roediger (Note 1) has noted that response withholding could be responsible for a variant of directed forgetting, posthypnotic amnesia.

One study that has addressed the question of whether subjects can actively forget some bits of encoded information was carried out by Timmins (1976, Experiment 6). His experiment was modeled on what may be referred to as the "harried short-order cook" phenomenon. Once an order has been filled, the cook must actively forget that order so that it will not interfere with his memory for unfilled orders. In Timmin's experiment, subjects were presented a list of words, some of which were repeated in the list. Only nonrepeated words were to be remembered. The initial presentation and encoding of a word was viewed to be analogous to the taking of an order, and its repetition to be analogous to the filling of that order, at which time it could (and should) be forgotten. Nonrepeated words represented orders that were unfilled and therefore were to be remembered. The subjects were told that when a word was repeated, they should "stop trying to remember it and forget it if you can." They were further told that it was to their advantage to actively forget the repeated items during the presentation of the list so that these items would not interfere with their memory for the nonrepeated items.

The results were consistent with the hypothesis of active forgetting. In a condition where no mention was made of forgetting the repeated items, the repeated items were remembered more often than the nonrepeated items (43% vs. 27%). However, with the active-forgetting instructions, the repeated (F) items were recalled slightly less often than were the nonrepeated items (21% vs. 27%), even though the subjects were told at test to try to recall all the items. This result, which is reminiscent of the Zeigarnik effect where uncompleted tasks are more likely to be remembered than completed tasks (filled orders), provides some support for the notion that subjects can actively block access routes to previously encoded information. However, as was noted above, we cannot be sure to what extent the subjects were engaging in active response withholding (suppression) at test, even though they were told to try to recall everything.

If it is the case that subjects initiate an inhibition process when a cue to forget is given, then a later countercommand making the F material R material may serve to free many of the blocked access routes. This result is typical in studies of posthypnotic amnesia (Evans & Kihlstrom, 1973; Freud, 1920/1952, p. 288; Kihlstrom & Evans, 1976) in which hypnotized subjects are given a suggestion to forget certain events with an accompanying countercommand that will later signal the removal of this suggestion. Such a countercommand could not negate an initial encoding deficit resulting from the F cue; but the countercommand does, in large part, reverse the previous recall impairment. An analogous experiment on nonhypnotic forgetting was conducted by Reed (1970). In his experiment, some letter trigrams were initially cued R, whereas others were initially cued F. On a later repetition of the trigrams, the F commands were reversed. Consistent with the inhibition release idea and at first glance inconsistent with the encoding deficit notion, subjects were able to recall just as many trigrams that were initially F cued as were initially R cued. Unfortunately, independent research on the repetition effect suggests that the magnitude of the effect may be greater if the initial presentation is more difficult to remember at the time of the second presentation (cf. the argument in Bjork, 1972, p. 228). Thus, Reed's results are not entirely persuasive because an initial encoding deficit owing to the F cues, rather than retrieval inhibition, could have been offset by enhanced encoding at repetition, rather than inhibition release.

Experiment 1

It is apparent, then, that new experimental paradigms are needed to circumvent past methodological limitations if retrieval inhibition is to be convincing as a contributor to directed-forgetting phenomena. The present paradigm involves a typical directed-forgetting procedure in which an unpredictable F cue is presented in the middle of certain lists. The innovation is that some material throughout the list (both precue and postcue) is learned incidentally, via a pleasantness-judging task, and is therefore not explicitly subject to the F cue. Thus, any effect of the F cue on the retention of the incidental items, for which selective rehearsal or response withholding would be unlikely, can be compared with the effect of the F cue on the to-be-learned items. If the F cue were to have no effect on the recall of the items that the subjects had no intent to rehearse for later recall (the to-be-judged words), then selective rehearsal would provide an adequate accounting of the results. If, on the other hand, the F cue were to affect the recall of the items learned incidentally as well as those learned intentionally, then disrupted retrieval for events occurring prior to the cue would provide a more plausible explanation of the results. Response withholding should not be a factor because the subjects are not told to forget the items learned incidentally, and hence there is no conflict created.

Method

Subjects. The subjects were 64 undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles. The subjects were tested in groups of four to seven. Course credit was given in exchange for participation in the experiment.

Materials and procedure. A list of 48 four-letter nouns was presented to subjects auditorily, with 7-sec interword intervals. To-be-learned and to-be-judged words were presented in strict alternation throughout the list. To-be-learned words were each preceded by an instruction to learn that word in preparation for a recall test at the end of the experiment (e.g., "learn *hand*"). The remaining words were each preceded by an instruction to judge that word on the basis of pleasantness (e.g., "judge *boat*"). The experiment was thus presented as a dual-requirement task: The subjects were told explicitly that the judged words were not to be learned. Specifically, the instructions stated that two studies were being conducted simultaneously to save time, one to see how memorable certain words were and one to see how pleasant certain words were.

At the midpoint in the list, half the subjects were given an instruction to forget the preceding to-be-learned words. They were told: "What you have done thus far has been practice; therefore, you should forget about all of the to-be-learned words that you have heard." The remaining subjects participated in the R-cue condition and were told: "The first portion of the list has now been presented; continue to try to remember the to-be-learned words that you have heard."

At the end of the list, all subjects were given a 3-min. distractor task, consisting of mathematical deductive reasoning problems. Then, half of the subjects in each cue condition (remember or forget) were told to try to recall all of the words that they had heard in the experiment, whereas the remaining subjects were given a word-recognition test. Four columns were provided for recall: one for the to-be judged words from the first half of the list, one for the to-be-judged words from the second half of the list, one for the to-be-learned words from the first half of the list, and one for the to-be-learned words from the second half of the list. In the F-cue condition, subjects were told that the "practice" items comprised the first half of the list and that it was important to write down these words as well as the words from the second half of the list. Eight min. were allowed for recall.

For the 32 subjects who were given the recognition test instead of the recall test, a sheet of paper was provided with 96 four-letter nouns typed on it, consisting of the 48 list items plus 48 distractors. The subject's task was to respond "yes" or "no" to each word depending on whether the word had appeared in the experimental list.

Results and Discussion

Word recall. With respect to the recall of the learn words, the typical directed-forgetting phenomena were obtained. Recall of the learn words from the first half of the list was poorer with the F cue than with the R cue, whereas recall of the learn words from the second half of the list was greater with the F cue. The results of the recall task are presented as a function of input serial position in Figure 1 and collapsed across serial position in Figure 2. Inspection of either figure shows that the recall of the judge words followed the same pattern as with the learn words, just at a lower level of performance overall. The Cue (remember or forget) × List Half interaction effect was significant, $F(1, 30) = 47.9$, $MS_e = .02$, $p < .001$, and analyses

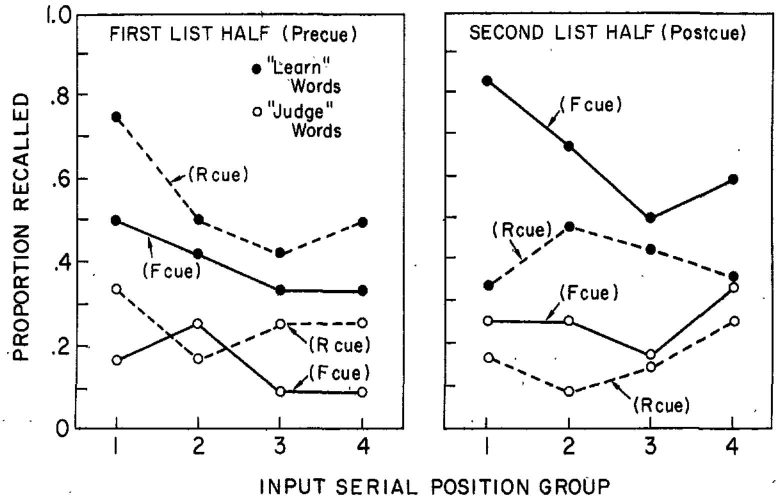

Figure 1. Proportion correct recall in Experiment 1 as a function of list half, type of midlist cue, and serial position group. (R = remember; F = forget.)

of the simple main effects showed that the above-mentioned relationships were significant for the judge words ($ps < .05$) as well as for the learn words ($ps < .001$). However, the Word-Type (learn vs. judge) × Cue × List Half interaction effect was also significant, $F(1, 30) = 8.3$, $MS_e = .02$, $p < .01$. The effect of the F cue on the recall of the judge words was not as great as the effect of the F cue on the recall of the learn words.

These results are consistent with the hypothesis that an F cue serves to initiate a process that inhibits the accessibility of a space of time in episodic memory. That is, not only were the to-be-learned items more poorly recalled following the F cue but so were items that the subjects had no intent to remember anyway. It is, therefore, difficult to explain these results in terms of a selective-rehearsal mechanism or in terms of response withholding. In a postexperimental debriefing session, none of the recall subjects and only three of the recognition subjects reported having suspected a memory test on the judge words.

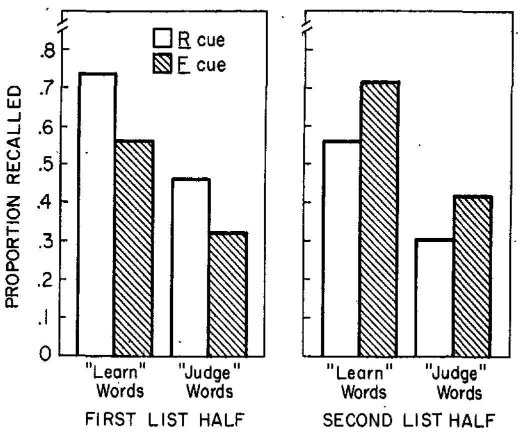

Figure 2. Proportion correct recall in Experiment 1 collapsed across serial position. (R = remember; F = forget.)

The greater effect of the F cue on the learn words is understandable since after the F cue, rehearsal processes would be devoted entirely to the second list-half learn words. The first list-half learn words would therefore receive less rehearsal than in the R-cue condition, and the second list-half learn words would receive more rehearsal. The judge words, on the other hand, are not rehearsed in any condition so such selective rehearsal arguments cannot apply to those items. In fact, the enhanced recall of postcue judge words in the F-cue condition is as surprising as the poorer recall of the precue judge words. Two sources of evidence imply that the subjects did not rehearse the judge words along with the learn words. First, on a postexperimental questionnaire, none of the 32 subjects reported that they suspected a "surprise" test on the judge words. Second, inspection of Figure 1 shows that there was no second list-half primacy effect for the judge items, whereas there is a strong second list-half primacy effect for the learn words following the F cue.

Finally, it is apparent from Figure 2 that more learn words were recalled than judge words, $F(1, 39) = 25.2$, $MS_e = .03$, $p < .001$. This outcome is intriguing because Hyde and Jenkins (1969, 1973) and others have shown that intent to remember has no greater effect on recall than certain semantic orienting tasks, such as the pleasantness judgment task in the present experiment. However, intent to learn has been found to yield greater delayed recall when intent is manifested in differential interitem organization (Battig & Bellezza, 1979). Given that the judge words were not rehearsed, it appears that they do not profit from the kind of interitem associations that are formed via cumulative rehearsal in the learn word case. Alternatively, perhaps the interassociation network for the learn words served to inhibit the retrieval of the judge words (Anderson & Bower, 1973).

Word classification and input–output order analyses. The obvious initial explanation of the above recall results would be that differential grouping of the learn and judge words in memory was rather poor; consequently, some judge words became, functionally, learn words for purposes of rehearsal. However, in the subjects' recall protocols, the discriminability of learn versus judge words was excellent. Less than 5% of the total words recalled of either type were misclassified with respect to the learn–judge distinction. A cell-by-cell breakdown of these results is presented in Table 1. On the basis of this finding, the hypothesis that judge words received inadvertent rehearsal because they were confused with learn words seems unlikely. Nonetheless, Experiment 2 was designed to provide a further, more conclusive test of such a possibility.

It was also observed that the subjects had relatively poor memory for the list-half membership of both learn and judge words presented before the F cue (see Table 1). The Cue (remember or forget) × List Half interaction effect was significant, $F(1, 30) = 11.54$, $MS_e = .07$, $p < .001$. Further, the order in which the F learn words were recalled was relatively unrelated to the order in which they were presented (see Table 2). These results are in contrast to those observed for words presented after the F cue, or before or after the R cue. This additional indication of disrupted retrieval of F items provides an interesting parallel with results from studies concerning posthypnotic amnesia (see Kihlstrom, 1977, for a review). Evans and Kihlstrom (1973), for example, reported that for subjects who are susceptible to hypnosis, events that can be remembered in spite of a hypnotic suggestion to forget typically are recalled in an order unrelated to the order of input. In contrast, a significant positive relationship is usually observed between input order and output order in the recall protocols of subjects who are relatively unsusceptible to hypnosis.

Table 1
Proportion Correct List-Half Classification Given Recall and Proportion Correct Learn–Judge Classification Given Recall

	List-half words			
	First		Second	
Word type and midlist cue	List-half classification	Learn–judge classification	List-half classification	Learn–judge classification
Intentional learn words				
Remember	.92	.92	.95	.95
Forget	.63	.95	.94	1.00
Incidental judge words				
Remember	.72	1.00	.69	.88
Forget	.36	.86	.75	1.00

Table 2
Average Rho Correlation Between Input and Output Order for Learn Words Recalled as Learn Words

	List-half words			
	First		Second	
Midlist cue	ρ	n	ρ	n
Remember	.72*	12	.70*	10
Forget	.29	8	.84*	8

Note. Correlation computed for those subjects who recalled three or more learn words.
* $p < .01$.

Thus, the hypothesis that directed forgetting involves disrupted retrieval provides a link with the literature on posthypnotic amnesia, suggesting that at least one of the underlying mechanisms is similar in the two situations in spite of the dramatic procedural differences. The inability of subjects to classify the F items as first list-half words is also consistent with the conclusion of Stern (1981) that human amnesia can, in some cases, be attributed to a context retrieval deficit.

It is instructive to note that disorganized retrieval in the present data was not simply a property of the items that were least memorable. As can be seen by comparing Figure 2 with Tables 1 and 2, the words presented in the second list half following the R cue also were recalled at a low level, comparable to that of the pre-F-cue words, but the classification performance was much more accurate for the post-R-cue words (.95 vs. .63 for the learn words and .69 vs. .36 for the judge words), and the input–output order correlation was highly significant for the post-R-cue words (.70). These results are consistent with findings reported within the hypnosis domain. Schwartz (1980), for example, found the retention of order information to be unrelated to the number of items recalled under hypnosis (r = .01).

One final result of interest from Table 1 is that the learn words were classified correctly more often than the judge words, $F(1, 30) = 67.8$, $MS_e = .03$, $p < .001$. This difference is consistent with the hypothesis that order information includes both temporal and relation-chaining cues (Lee & Estes, 1977). With the judge words, only temporal cues would be available since cross-item interassociations were not intentionally developed. This result provides additional indirect evidence that the subjects did not rehearse the judge words during the presentation of the list.

Word recognition. The results of the yes–no word-recognition test are presented in Figure 3. An analysis of variance (ANOVA) revealed no significant differences among the means. This outcome is consistent with the hypothesis considered above that the differences observed in the recall results are largely retrieval phenomena. The complete absence of significant effects of list half, item type,

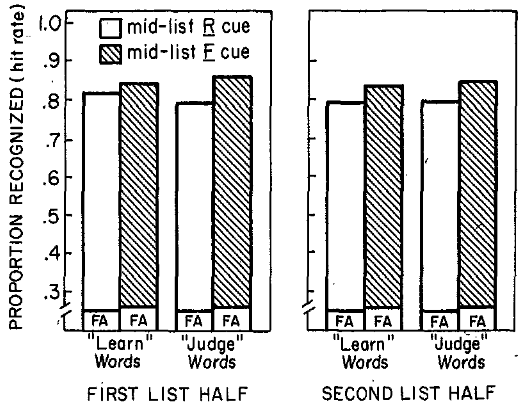

Figure 3. Proportion correct recognition in Experiment 1 as a function of list half and type of midlist cue (R = remember; F = forget. FA = false-alarm rate.)

and cue type is quite remarkable in view of the dramatic effects of those same variables reflected in the various measures of recall performance reported above. The F cue clearly had its effects primarily on the retrieval side.

Alternative hypotheses. There are two plausible explanations for the pattern of recall results described above other than the retrieval-inhibition hypothesis. First, perhaps the judge words became unavoidably interassociated with the learn words to some extent. The effect of the F cue on the judge words could then be explained as a disrupted-rehearsal phenomenon without reference to retrieval processes. Two aspects of the data from Experiment 1 argue against this hypothesis; namely, the judge words were not confused with the learn words in the classification data, and there was no second list-half primacy effect for the judge words following the F cue. Nevertheless, Experiment 2 was designed to render the learn and judge words completely nonconfusable with each other; they were drawn from two distinct semantic categories, nouns representing animate entities and nouns representing inanimate entities.

The second alternative hypothesis centered on the possibility that the results of Experiment 1 can be attributed to differential output interference. Perhaps with an F cue, the subjects recalled the words from the second half of the list first, producing enhanced recall of those words and depressed recall of both the learn and judge words from the first list half. Experiment 3 and 4 tested whether the results of Experiment 1 can be attributed to such differential patterns of output interference rather than to retrieval inhibition.

Experiment 2

As mentioned above, if some of the judge words were inadvertently included in the subjects' rehearsal sets, then perhaps the effect of an F cue on the precue judge words is attributable to the termination of rehearsal of those words as well as the precue learn words. This hypothesis was evaluated in Experiment 2 by selecting the judge words from a semantic category that is clearly different from that of the learn words and by informing the subjects of this difference before the experiment began. Experiment 2 also provided an opportunity to replicate the pattern of recall results from Experiment 1 with the learn–judge distinction more clearly defined.

Method

Subjects. The subjects were 46 undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles. The experiment was carried out with groups of 4 to 7 subjects per session. Course credit was given in exchange for their participation. Twenty-one subjects were assigned to the F-cue condition, and 25 to the R-cue condition.

Materials. Forty words were selected for use in Experiment 2, 20 nouns representing animate entities such as painter and 20 nouns representing inanimate entities such as magazine. Words with high meaningfulness values (m = 5.04 to 7.96) were selected and the two sets of 20 words were matched in a pairwise manner on meaningfulness using the Paivio, Yuille, and Madigan (1968) norms. As in Experiment 1, two tape recordings of the 40 words were made, one with the learn words as animate entities and one with the judge words as animate entities.

Procedure. Only two modifications of the procedure from Experiment 1 were made. Just prior to list presentation, the subjects were told:

> To help you keep the words that you are supposed to learn separate from the words that you are supposed to judge, we have constructed the list such that all of the learn (judge) words are inanimate objects whereas all of the judge (learn) words are animate entities.

The second modification of the procedure was to shorten the list to 40 words. The list was shortened in an attempt to raise recall performance for the judge words so as to evaluate more accurately the effect of the F cue on those items. The method for testing recall of the words was identical to that of Experiment 1, but recognition performance was not evaluated in this experiment.

Results and Discussion

Word recall. The ANOVA performed on the word-recall data showed that, as in Experiment 1, more learn words than judge words were recalled, $F(1, 44) = 55.1$, $MS_e = .06$, $p < .001$. In addition, the Cue (remember or forget) × List Half interaction effect was significant, $F(1, 44) = 53.4$, $MS_e = .03$, $p < .001$. Inspection of Figure 4 shows that more words were recalled from the first list half in the R-cue condition ($p < .05$), whereas more words were recalled from the second list half in the F-cue condition ($p < .05$). Most important, this pattern held for the judge words as well as for the learn words, as the Cue × List Half × Word Type (learn vs. judge) interaction effect was not significant, $F(1, 44) = 3.59$, $MS_e = .04$, $p > .05$. Thus, with the judge words distinguished from the learn words by semantic category (animate vs. inanimate), the effect of the F cue on the recall of the precue judge words was still evident. This result suggests that the pattern of recall obtained in Experiment 1 is robust and that the hypothesis in question, namely, that the judge words in Experiment 1 were inadvertently intertwined with the learn words during rehearsal, is unlikely.

Word classification and input–output order. Because the learn and judge words were distinguished by semantic category, only the list-half classification errors were of interest here. The classification data are summarized

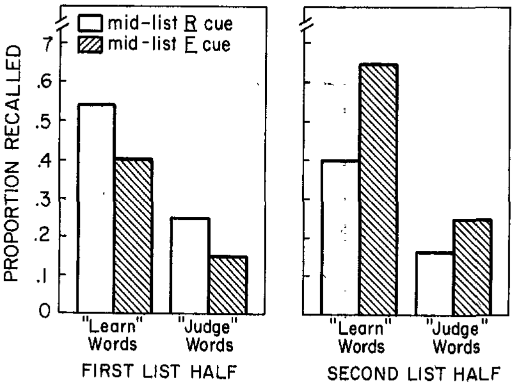

Figure 4. Proportion correct recall in Experiment 2. (R = remember; F = forget.)

Table 3
Proportion Correct List-Half Classification Given Recall

Word type and midlist cue	List-half words: First	Second
Intentional learn words		
Remember	.96	.90
Forget	.89	.97
Incidental judge words		
Remember	.84	.73
Forget	.72	.89

in Table 3. As in Experiment 1, the learn words that were recalled were classified correctly by list half more often than were the judge words that were recalled (93% vs. 79% correct), $F(1, 44) = 29.5$, $MS_e = .04$, $p < .001$. Also as in Experiment 1, the Cue (remember or forget) × List Half interaction effect was significant, $F(1, 44) = 7.19$, $MS_e = .06$, $p < .01$. With the R cue, list-half classifications were more often correct for first list-half words than for second list-half words (90% vs. 82% correct, $p < .05$); whereas with the F cue, list-half classifications were more often correct for second list-half words (93% vs. 81% correct, $p < .01$). Thus, as in Experiment 1, the subjects clearly had difficulty classifying the words that were presented prior to the F cue. Again, this result is consistent with the hypothesis that a cue to forget affects recall, in part, through disruption of retrieval processes.

Table 4
Average Rho Correlation Between Input and Output Order for Words Classified Correctly

Word type and midlist cue	First: ρ	First: n	Second: ρ	Second: n
Intentional learn words				
Remember	.74**	23	.76**	20
Forget	.32*	17	.56**	20
Incidental judge words				
Remember	.62**	16	.59**	11
Forget	.05	13	.50**	11

Note. Correlation computed for those subjects who recalled three or more words in that cell.
* $p < .05$. ** $p < .001$.

Additional evidence for disrupted retrieval is presented in Table 4. As in Experiment 1, the correlation between input order and output order was found to be depressed for the learn words presented prior to the F cue. Given the somewhat higher level of recall in this experiment, these correlations could also be examined for the judge words, and the same pattern was observed as for the learn words. In sum, these data replicate and extend the classification and input–output order data from Experiment 1 and further suggest that an F cue leads to disrupted retrieval.

Experiment 3

Experiment 3 was designed to ascertain whether differential output order in the F-cue and R-cue conditions could be responsible for the foregoing results. If, following an F cue, subjects choose to start their recall with the R-word list half (that is, the second list half), then recall of the second list half might be enhanced, and recall of the first list half might be depressed compared with the R-cue case. Presumably, subjects in the R-cue condition would show no list-half preference in starting their recall (or, possibly, would tend to start with the first list half). The design of Experiment 3 permitted measurement of the average output position for any given word in each condition.

Method

Subjects. The subjects were 32 undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles. They served in groups of 4 to 7 and received course credit for their participation. Sixteen subjects were assigned to the F-cue condition, and 16 to the R-cue condition.

Materials. The materials were identical to those used in Experiment 1 except the list was shortened to 40 words as in Experiment 2.

Procedure. The procedure was the same as that used in Experiment 1 except for the method of testing word recall. In this experiment, the subjects were instructed to record the words in one long column, independent of whether the word was a learn word or a judge word and independent of where in the list the word was presented. The subjects were told to write the words down in the order in which they came to mind. To avoid any interference with this recall procedure, the subjects were not required to classify the words in any way during recall.

Results and Discussion

Word recall. The word-recall results are presented in Figure 5. As in the previous two

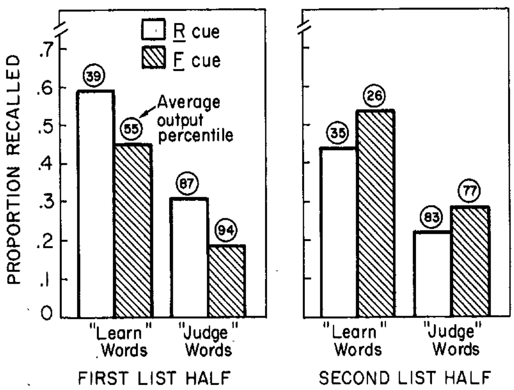

Figure 5. Proportion correct recall in Experiment 3. (R = remember; F = forget.)

experiments, more learn words were remembered than judge words, $F(1, 30) = 31.4$, $MS_e = .07, p < .001$, and the Cue (remember or forget) × List Half interaction was significant, $F(1, 30) = 29.2$, $MS_e = .07, p < .001$. The latter result reflects the depressed recall from the first list half in the F-cue condition ($p < .01$). Also as before, the Cue × List Half × Word Type (learn or judge) interaction effect did not reach significance, $F(1, 30) = 3.60$, $MS_e = .06$, $p > .05$, thus providing a second replication of the inhibitory effect of the F cue on the recall of the pre-F-cue judge words. The basic pattern of effects in Experiments 1, 2, and 3 (Figures 2, 4, and 5) is remarkably similar across studies.

Output order. The output-interference hypothesis predicts that with the F cue, the words from the second list half are recalled first, before the words from the first list half are recalled. To evaluate this prediction, a measure of average output position for each of four classifications of words was computed for each subject. The four classifications were first-half learn words, second-half learn words, first-half judge words, and second-half judge words. Because the nominal output position is confounded between subjects with differences in the number of words recalled, the specific measure of output position was an output-position percentile, as developed by Bjork and Whitten (1974). The output-position percentile for a given word that is recalled was computed as nominal output position divided by total number of words recalled by that subject then multiplied by 100.

The output-position percentiles averaged across subjects are included in Figure 5 just above each bar. A lower value indicates that words from that classification tended to be output earlier in recall. As can be seen, the learn words were recalled before the judge words, $F(1, 30) = 25.4$, $MS_e = .05, p < .001$, and with the F cue, the words from the second list half were recalled prior to the words from the first list half. The Cue × List Half interaction effect was significant, $F(1, 30) = 8.2$, $MS_e = .07$, $p < .01$, and this pattern was not significantly different as a function of word type (learn vs. judge). Thus, the output-interference hypothesis cannot be ruled out.

Input–output order. As was found in Experiments 1 and 2, the correlation between input presentation order and output recall order was found to be depressed for words presented prior to the F cue. The rho correlations from the present experiment are presented in Table 5. These data constitute the second replication of the effect of the F cue on the input–output order relationship.

Experiment 4

The results of Experiment 3 provided a second replication of the pattern of recall obtained in Experiment 1. The analysis of the output-position data, however, suggests that an output-interference explanation of the recall results cannot be ruled out. Thus, Experiment 4 was designed to provide a more

Table 5
Average Rho Correlation Between Input and Output Order for Words Classified Correctly

Word type and midlist cue	List-half words: First ρ	n	Second ρ	n
Intentional learn words				
Remember	.75**	14	.72**	11
Forget	.34	11	.68**	12
Incidental judge words				
Remember	.67*	8	.64*	7
Forget	.22	6	.59*	7

Note. Correlation computed for those subjects who recalled three or more words in that cell.
* $p < .05$. ** $p < .01$.

direct test of the output-interference hypothesis. In this experiment, the order of recall of the first list half versus the second list half was controlled between subjects. According to the output-interference notion, controlling the order of list-half output should eliminate the effects of the F cue on the judge words.

Method

Subjects. The subjects were 92 undergraduates from the introductory psychology course at the University of California, Los Angeles. Each subject was randomly assigned to one of two list conditions (R cue vs. F cue) and to one of two test conditions (recall the first vs. the second list half first). The subjects were tested in groups of 4 to 7, with 43 subjects participating in the R-cue condition and 49 participating in the F-cue condition.

Materials and procedure. The materials were the same as in Experiment 3. The procedure differed from that in the previous experiments by the nature of the recall test. In the present experiment, some of the subjects were required to write down the words from the first list half first, whereas the remaining subjects were required to write down the words from the second list half first. Three min. were allowed for recall from each list half, and these two recall tasks were carried out on separate sheets of paper. The subjects were not asked to classify the words by word type (learn vs. judge) in this experiment.

Results and Discussion

Word recall. The word-recall results are presented in Figure 6. An ANOVA confirmed that these data represent a third replication of the basic pattern of recall found in Experiment 1. More learn words were recalled than judge words, $F(1, 88) = 53.4$, $MS_e = .06$, $p < .001$. With the R cue, more words were recalled from the first list half, whereas with the F cue, more words were recalled from the second list half; the Cue × List Half interaction effect was significant, $F(1, 88) = 86.9$, $MS_e = .01$, $p < .001$. This pattern held for both the learn words and the judge words.

Most important here, the forced order of output (first list-half output first vs. second list-half output first) did not interact significantly with any other factor in the design (all Fs < 1.07). This result rules out an output-interference explanation of the effects of the F cue in the present experiments.

Word classification. The subjects in this experiment were not asked to classify the words that they recalled on the basis of the learn–judge distinction. However, classification errors could be examined with respect to list half as in Experiments 1 and 2. These data are presented in Table 6. It will be recalled from Experiment 3 that the subjects tended to recall the second list-half items first in the F-cue condition. As can be seen in Table 6, when the subjects were asked to recall the second list-half words first, the words that were classified correctly least often were those presented before the F cue, both learn and judge, just as in Experiments 1 and 2. However, when the subjects were asked to recall the first list-half words first, an effect of the output-order manipulation was evident, as the Cue × List Half × Output Order interaction effect was significant, $F(1, 88) = 4.01$, $MS_e = .06$, $p < .05$. Specifically, with the F cue, the first list-half words were classified correctly more often when the first list-

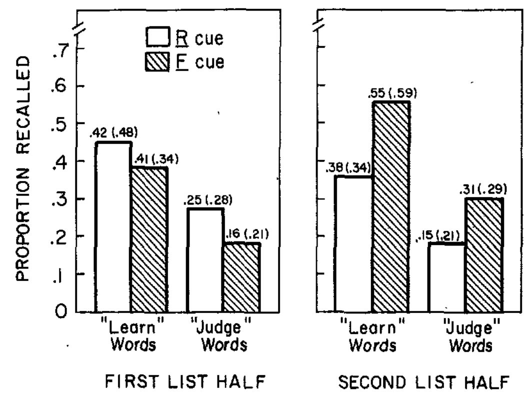

Figure 6. Proportion correct recall in Experiment 4. (R = remember; F = forget.)

Table 6
Proportion Correct List-Half Classification Given Recall

	List-half words			
	First		Second	
Word type and midlist cue	Output last	Output first	Output first	Output last
Intentional learn words				
Remember	.92	.94	.95	.82
Forget	.70	.81	.96	.95
Incidental judge words				
Remember	.91	.90	.99	.69
Forget	.52	.82	.97	.93

half words were output first; and with the R cue, the second list-half words were misclassified more often when the first list-half words were output first.

Both of these outcomes are interpretable if subjects tended to write down words they were unsure about on the first page of testing. It is not clear, however, why the same outcome did not obtain for the remaining two cases. With the F cue, the second list-half words were not misclassified more when the first list-half words were output first; with the R cue, the first list-half words were not classified correctly more often when the first list-half words were output first. Perhaps the subjects in these conditions were less likely to write down words they were unsure about on the first page because a greater number of words were accessible to them for recall in those cases. That is, there may have been a reduced tendency to "pad" the recall on the first page. Because of these potential confounding biases, input–output order correlations were not computed for Experiment 4.

General Discussion

In four experiments, a midlist cue to forget the words in the first half of a list served to depress later recall of those words. This effect was observed for items that the subjects had no a priori intent to remember (items learned incidentally) as well as for R words. It does not seem possible to explain the present results without postulating that a cue to forget has the power to initiate a process that blocks or inhibits access routes to the F items. If the F cue served only to terminate rehearsal of the pre-F-cue learn words, then the forget instruction should have had no effect on the recall of the words learned incidentally, which were not rehearsed in any case. Because recognition performance was unimpaired, it appears that the inhibition induced via the F cue took the form of retrieval blocking. Two less dramatic, less interesting interpretations of the depressed recall of the pre-F-cue incidental items were ruled out, one by Experiments 1 and 2 and the other by Experiments 3 and 4.

Two other aspects of the recall data suggest that disrupted retrieval plays a significant role in intentional-forgetting experiments of the type presented here. First, the subjects could not remember whether the words that were presented prior to the F cue were presented in the first half of the list or in the second half of the list. That is, they exhibited list-half source amnesia for words presented prior to the F cue. Second, the average correlation between input order and output order for words presented prior to the F cue was greatly depressed. These results are comparable to those reported from studies of posthypnotic amnesia (Evans & Kihlstrom, 1973; Kihlstrom & Evans, 1976), in which disrupted retrieval also has been implicated. Thus, intentional forgetting and posthypnotic amnesia appear to have strong parallels such that future work in one discipline should impact on the other. For example, it has been shown that subjects who experience posthypnotic amnesia do not have differentially poor waking memories (Kihlstrom & Twersky, 1978), but future research may reveal that these individuals are more likely to forget when the forgetting is intentional, as in the present experiments. That is, there may be a subset of individuals who are extremely efficient at forgetting. Consistent with this argument is the recent finding by Geiselman et al. (in press) that subjects who exhibit low F-item recall in a nonhypnotic directed-forgetting experiment are the subjects who are most likely to exhibit low recall (prior to the countercommand) in a hypnotic amnesia experiment.

The comparison between paradigms must be tempered, however, because it appears that a hypnotic suggestion to forget can have a profound effect on the accessibility of episodic information that is well learned (Kihlstrom, 1980), whereas under typical laboratory conditions, a nonhypnotic cue to forget has little effect on memory for items that receive elaborate precue processing (Bugelski, 1970). Further, once a hypnotic suggestion to forget is lifted, previously hypnotized subjects can retrieve much of the target information from memory. No convincing procedure for reversing a nonhypnotic forget command has been reported. Thus, it is apparent that procedures that rely on hypnosis obtain greater control over the subjects' recall performance, whereas the nonhypnotic procedures from cognitive psychology lead to differential encoding of the items in addition to disrupted retrieval.

Finally, a concluding speculation may be in order on the adaptive character of retrieval blocking as an updating mechanism (see Bjork, 1976). Any ongoing information-processing system needs some mechanism for updating information, that is, some means by which to prevent old, out-of-date information from interfering with new, current information. In a sense, computers are the ultimate in efficiency in updating because storing new information at a given memory location obliterates the prior entry. However, should it become important to gain access to the out-of-date information, such an updating system is not optimal. In comparison, inhibition that takes the form of retrieval blocking has some desirable properties as an updating mechanism. On the one hand, information that is not retrievable is also noninterfering. On the other hand, such "blocked" information is still readily recognizable as having occurred before, and it appears that upon becoming pertinent again (that is, changing from F to R), such information shows a repetition effect (Reed, 1970).

Reference Note

1. Roediger, H. L. Directed forgetting and post-hypnotic amnesia. Paper presented at the meeting of the American Psychological Association, Toronto, August 1978.

References

Anderson, J. R., & Bower, G. H. *Human associative memory.* Washington, D.C.: Winston & Sons, 1973.

Battig, W. F., & Bellezza, F. S. Organization and levels of processing. In C. R. Puff (Ed.), *Memory organization and structure.* New York: Academic Press, 1979.

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

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

Bjork, R. A. The updating of human memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 12). New York: Academic Press, 1976.

Bjork, R. A., & Geiselman, R. E. Constituent processes in the differentiation of items in memory. *Journal of Experimental Psychology: Human Learning and Memory,* 1978, *4,* 347–361.

Bjork, R. A., & Whitten, W. B. Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology,* 1974, *6,* 173–189.

Block, R. A. Effects of instructions to forget in short-term memory. *Journal of Experimental Psychology,* 1971, *89,* 1–9.

Bugelski, B. R. Words and things and images. *American Psychologist,* 1970, *25,* 1002–1012.

Davis, J. C., & Okada, R. Recognition and recall of positively forgotten items. *Journal of Experimental Psychology,* 1971, *89,* 181–186.

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

Evans, F. J., & Kihlstrom, J. Posthypnotic amnesia as disrupted retrieval. *Journal of Abnormal Psychology,* 1973, *82,* 317–323.

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

Geiselman, R. E. Positive forgetting of sentence material. *Memory & Cognition,* 1974, *2,* 677–682.

Geiselman, R. E. Effects of sentence ordering on thematic decisions to remember and forget prose. *Memory & Cognition,* 1977, *5,* 323–330.

Geiselman, R. E., et al. Mechanisms of hypnotic and non-hypnotic directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* in press.

Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology,* 1969, *82,* 472–481.

Hyde, T. S., & Jenkins, J. J. Recall of words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 471–480.

James, W. *The principles of psychology* (Vol. 2). New York: Holt, 1890.

Johnson, D. A. Pupillary responses during a short-term memory task: Cognitive processing, arousal, or both? *Journal of Experimental Psychology,* 1971, *90,* 311–318.

Jongeward, R. H., Woodward, A. E., & Bjork, R. A. The relative roles of input and output mechanisms in directed forgetting. *Memory & Cognition,* 1975, *3,* 51–57.

Kihlstrom, J. F. Models of posthypnotic amnesia. In W. E. Edmonston (Ed.), *Conceptual and investigative approaches to hypnosis and hypnotic phenomena* (Vol. 296). New York: New York Academy of Sciences, 1977.

Kihlstrom, J. F. Posthypnotic amnesia for recently learned material: Interactions with "episodic" and "semantic" memory. *Cognitive Psychology,* 1980, *12,* 227–251.

Kihlstrom, J. F., & Evans, F. J. Recovery of memory after posthypnotic amnesia. *Journal of Abnormal Psychology,* 1976, *85,* 564–569.

Kihlstrom, J. F., & Twersky, M. Relationship of posthypnotic amnesia to waking memory performance. *The International Journal of Clinical and Experimental Hypnosis,* 1978, *26,* 292–306.

Lee, C. L., & Estes, W. K. Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 395–418.

Paivio, A., Yuille, J. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph*, 1968, *76*, (1, Pt. 2).

Reed, H. Studies of the interference processes in short-term memory. *Journal of Experimental Psychology*, 1970, *84*, 452–457.

Reitman, W., Malin, J. T., Bjork, R. A., & Higman, B. Strategy control and directed forgetting. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 140–149.

Ribot, T. A. *Diseases of memory: An essay in the positive psychology*. New York: Appleton-Century-Crofts, 1882.

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

Roth, D. M. *Roth memory course*. Santa Monica, Calif.: Motivation, 1961. (Originally published, 1918.)

Schwartz, W. Hypnosis and episodic memory. *The International Journal of Clinical and Experimental Hypnosis*, 1980, *28*, 375–385.

Stern, L. D. A review of theories of human amnesia. *Memory & Cognition*, 1981, *9*, 247–262.

Timmins, W. K. Strategies of remembering: A study of directed forgetting. Unpublished doctoral dissertation, Australian National University, Canberra, 1976.

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

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

Woodward, A. E., Bjork, R. A., & Jongeward, R. H. Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 608–617.

Received April 19, 1982
Revision received September 13, 1982 ■