FORGETTING AND REMEMBERING IN FREE RECALL: INTENTIONAL AND UNINTENTIONAL

ADDISON E. WOODWARD, JR. AND ROBERT A. BJORK

Albion College
University of Michigan

Two free recall experiments were designed to study the processes by which Ss remember some items and forget others. In both experiments, Ss were cued immediately after each word in a list whether to remember (R word) or to forget (F word) that word. After each of six such lists, Ss were asked to recall the R words and to avoid recalling the F words; in general, Ss were remarkably able to do both. At the end of the experiment, Ss were asked, without forewarning, to recall any and all R words and F words they could remember. In Exp. I, final recall of F words was very poor: they seemed not to be in memory. In Exp. II, which employed categorized lists, Ss recalled F words quite well given that there were R words in the same semantic category. The results suggest that the differential rehearsal devoted to R words operates primarily on retrieval rather than on storage.

Our attempts to understand human memory as exhibited in the free recall paradigm tend to concentrate on the items S recalls. We concern ourselves primarily with characteristics of S's successful item retrievals, for example, output order, measures of subjective organization and clustering over trials, and probability of recall as a function of serial position, list length, and other variables. Failure to recall, however, is as much a property of memory as is successful recall; we cannot understand remembering independent of forgetting.

It is understandable that we look primarily at recall rather than nonrecall in the free recall paradigm. As Bjork (1970b) has pointed out, standard memory paradigms study intentional learning and, hence, incidental forgetting. From S's standpoint, however, the typical free recall task easily exceeds his capacity to perform perfectly, and he must sacrifice some items for the sake of other items. He must, in a sense, choose what to remember and, thereby, what to forget. In choosing to rehearse, relate, and group some items, he drastically reduces the likelihood that he will recall other items.

This paper reports two experiments designed to assess intentional forgetting and incidental remembering of to-be-forgotten words in a free recall list as well as performance on typical to-be-remembered items. Both experiments involved cues to Ss to forget individual words in a free recall list. Similar cueing-to-forget procedures have been employed in a number of recent studies. Some of these studies have attempted to clarify interference mechanisms in memory (Bjork, 1970b; Bjork, LaBerge, & Legrande, 1968; Elmes, 1969a, 1969b; Elmes, Adams, & Roediger, 1970; Reed, 1970; Turvey & Wittlinger, 1969); others have been concerned with motivational factors in memory (Weiner & Reed, 1969), causes of the primacy effect in free recall (Bruce & Papay, 1970), and other problems. The intentional-forgetting paradigm represents a tool with which to attack a variety of problems in the study of memory.

In general, intentional forgetting experiments involve a signal to S that he can forget the items presented prior to the signal. In many cases, the effects of such
an instruction are quite remarkable. Under some circumstances, for example, a forget instruction eliminates all proactive interference owing to to-be-forgotten items on the recall of to-be-remembered items. In all past experiments, to-be-remembered items and to-be-forgotten items have been blocked in time. Whether such temporal blocking of forget items and remember items might be an important factor (possibly a necessary condition) in eliminating proactive interference owing to forget items was a main concern of the present research. In the two experiments reported here, Ss were signaled word by word during the presentation of a free recall list whether to remember (R word) or to forget (F word) each individual word.

A second innovation was designed to judge the extent to which F items are stored in long-term memory. What one would like to do, of course, is instruct Ss to forget items and then, at the end of some retention interval, test for Ss' retention of those items. In the face of such a contradictory procedure, however, Ss are unlikely to consistently forget when told to do so, or, for that matter, to recall when told to do so. Weiner and Reed (1969) have worked with this procedure in spite of its problems, but there seems no way to entirely avoid the difficulties inherent in cueing Ss to forget items they know they will later have to recall. In the present experiments, Ss were presented a series of lists and were required to recall the R words from a given list immediately after the list was presented. At the end of the experiment, after all of the lists had been presented and recalled, Ss were asked, without forewarning, to recall any and all words they could remember from the lists they had seen, independent of whether those words had been R words or F words.

**Experiment I**

**Method**

**Subjects.**—The Ss were 36 undergraduates at the University of Michigan. They were paid $1.00 plus any bonuses that accrued from the payoff system employed in the experiment.

**Materials and apparatus.**—Every S viewed six 24-word lists. The words were all common four-letter nouns. After each list there was a recall period, and after all six lists were presented and recalled there was a final recall period for all items from all lists.

The apparatus was a high-speed (change time less than .05 sec.) memory drum. The words, the cues to forget or to remember, and the instructions to recall or get ready for the next trial all appeared in the same window. The timing of advances of the memory drum was controlled by a high-speed paper-tape reader reading a prepunched tape.

**Design.**—After each successive word in a list, a colored (red or green) dot appeared as a cue to S whether to forget or to remember that word. For half of the Ss, a green dot meant remember and a red dot meant forget, and, for the other Ss, a red dot meant remember and a green dot meant forget.

An individual word in a list was presented for 1, 2, or 4 sec. The colored-dot cue was shown immediately thereafter for 1 sec. before the next word appeared. Suitable counterbalancing techniques were used to insure that, across Ss, every word was both an R word and an F word at each of the presentation times. Also, F and R words at each of the three presentation times appeared equally often at each serial position. Every quarter of a list (6 words) contained 3 R words and 3 F words, 1 at each presentation time. Hence, there were 12 R words and 12 F words in each list, 4 at each presentation time.

**Procedure.**—The Ss were run individually. Every S was read a set of instructions and was shown a practice list of 24 two-digit numbers to familiarize him with the procedure. After the practice list, Ss were informed that a payoff matrix would be in effect during the experiment: they received a 1¢ reward for each R word recalled following a list, and they lost 1¢ for each F word recalled.

Each of the six lists was preceded by a 1-sec. ready signal and followed by a 30-sec. recall period during which Ss wrote down any R words they could remember on a response sheet. After the recall of the sixth list, there was a phony debriefing period lasting several minutes, and, finally, Ss were asked to recall any and all words they could remember from any of the lists they had seen. They were told that they would receive a 1¢ bonus for any word they could recall, independent of whether the word was an F word or an R word. When S could recall no more words, he was asked to circle any word among those he had recalled that he thought was an F word.

**Results**

In Table 1 are shown the immediate and final recall probabilities of R words and F words in each of the six lists. Overall, in immediate recall, Ss were remarkably able to recall R words and to not recall F words. They simultaneously recalled 50% of the R words and intruded less than 2% of the F words. On the average, a total of fewer
than 1.5 F words appeared during the course of all six immediate recall efforts for any one S.

The immediate nonrecall of F words might be attributable to either (a) actual forgetting or (b) active suppression of those words during immediate recall. Comparing the immediate and final recall probabilities in Table 1 suggests that active suppression is a relatively minor factor. The fact that final F-word recall exceeds immediate F-word recall, even though the R-word data show a retention loss from immediate to final recall, indicates that suppression did occur, but in absolute terms the recovery of F words during the final recall is very small. An average S recalled only 3.4 of the 72 F words shown.

The recall probabilities in Table 1 are quite consistent across the six lists. In the final recall of R words there is a recency effect across the six lists, and the immediate recall of R words in List 1 is somewhat higher than in Lists 2–6, but the lack of any sizable changes with list number is more impressive than are those two effects. No trends at all are apparent in the recall of F words. All subsequent analyses in this section combine the recall data from the six individual lists.

In Table 2, immediate and final recall probabilities are shown as a function of presentation rate. The chief motivation for manipulating presentation time was to assess whether a forget signal would become progressively less effective the longer a word was shown before the signal. In particular, it seemed likely that there would be a greater recovery in the final recall of F words with increasing presentation time.

Although the immediate and final recall probabilities of F words do increase with presentation time, the increases are small in absolute terms. There is no clear relation at all between the recall of R words and presentation time. In general, the manipulation of presentation time was ineffectual.

It is easy to understand, after the fact, why recall probabilities were not heavily influenced by presentation time. The Ss reported that they tried not to do anything with a word during its presentation beyond keeping it in sight. They waited to see if the signal at the end of the presentation interval designated the current word as an R word before they tried to rehearse it or relate it to any other R words they could remember.

The immediate and final recall probabilities of R words and F words are shown in Fig. 1 as a function of serial position. Except for a somewhat larger primacy effect and a somewhat smaller recency
effect in immediate recall, the two serial position curves for R words replicate the findings of Bjork (1970a), Craik (1970), and Craik, Gardiner, and Watkins (1970). In particular, in those studies and in Fig. 1, the long-term retention of recency words is inferior to the long-term retention of words in the middle of the list, even though the immediate recall of recency words is much better than the immediate recall of words in the middle of the list.

The immediate and final serial position curves for F words do not show appreciable effects of either primacy or recency. The lack of any such effects argues against theories that attribute primacy or recency effects to input or output advantages intrinsic to the early or late positions in a list, respectively (for more detailed considerations of the issue, see Bjork, 1970a, and Bruce & Papay, 1970).

"Forget" and "remember" labeling.—After Ss had completed their final recall of any words they could remember from the preceding lists, they were asked to circle any words on their response sheet that they judged to be F words. On the average, each of the 36 Ss recalled 22.2 words on the final recall test. Of those 22.2 words, 16.8 were R words, 3.4 were F words, and 2.0 were intrusions. The Ss circled 3.3 words on the average, very close to the average number of F words actually recalled. They mistakenly circled only about 6% of the R words, and they correctly circled about half (45%) of the F words. Interestingly, they circled one-third of their intrusions.

Whether S labeled a given word in his final recall as an F word or not was heavily influenced by whether that word had been immediately recalled or not. In general, words recalled both immediately and finally tend to be judged as R words whether they were in fact R words or F words, and words recalled finally but not immediately are more likely than other words to be judged as F words, whether they were in fact F words or R words. In particular, the few F words (14) recalled both immediately and finally appear to be words that were mistakenly encoded as R words by the Ss: all but one of those words were judged to be an R word.

**Experiment II**

The results of Exp. I suggest that to-be-forgotten items are really "forgotten." Even when Ss were rewarded for recalling F words during the final recall, they were able to recall less than 5% of the F words they had been presented. It seems plausible that the few F words that were recalled arise from the infrequent occasions when an individual F word was either misinterpreted as an R word or had some very salient association or significance for a particular S.

It is also possible that F words exist in memory but are not accessible. There is some evidence (for example, Thomson & Tulving, 1970; Tulving & Pearlstone, 1966) that recall failures in standard free recall experiments are frequently the result of retrieval failures. If F words exist in memory at the time of the final recall, but are peculiarly nonretrievable, one would like to understand why. One possibility derives from a theory of intentional forgetting proposed by Bjork (1970b). The theory assumes that Ss are able to accomplish a kind of differential grouping of to-be-remembered items that functionally separates them from to-be-forgotten items. Thus, to-be-forgotten items do not provide interference during recall because they are segregated from to-be-remembered items; that is, to-be-forgotten items are not absent from memory, but, rather, are differentiated in some manner. The theory also assumes that during presentation, Ss selectively devote all rehearsal activity to to-be-remembered items.

It is not clear, however, in terms of Bjork's theory, why Ss in Exp. I were unable to retrieve F words when they were trying to do so. Experiment II was motivated by the desire to clarify the extent to which F words exist in memory and the mechanisms by which they become noninterfering and nonretrievable.

Experiment II was designed both to test whether Ss could divide the words in a
natural semantic cluster (for example, PINE, OAK, BIRCH, and MAPLE) into to-be-remembered and to-be-forgotten subsets when instructed to do so and to investigate whether such a semantic relationship between R words and F words would facilitate the retrieval of F words during the final recall.

**Method**

Experiment II involved the same general procedure as Exp. I: six individual lists were presented one by one, there was an immediate recall test after each list, and there was a delayed final recall test following all six lists. Presentation time was not varied, however, and each list was constructed from six 4-word semantic categories rather than from 24 unrelated words.

**Subjects.**—The Ss were 48 undergraduates at the University of Michigan. They were paid $1.00 plus any earned bonuses.

**Materials and apparatus.**—The six 4-word categories in each 24-word list were drawn from a recent revision of the Connecticut Category Norms (Battig & Montague, 1969). Every word in a given category was one of the 10 most frequent associates of that category. The same high-speed memory-drum apparatus used in Exp. I was used in Exp. II.

**Design.**—As in Exp. I, a colored dot (red or green) appeared for 1 sec. after each successive word as a cue to S to forget or to remember that word. Every word was presented for 2.3 sec., a time equal to the average presentation time in Exp. I. The four words in any one category were distributed throughout a list: one word from each of the six categories in a list appeared in every quarter of the list. The order of the category instances in any one quarter of a list was random except that across the entire list, at least two items from other categories intervened between successive items of the same category.

Every list contained one 0R-4F category (no R words, four F words), one 1R-3F category (one R word, three F words), two 2R-2F categories (two R words, two F words), one 3R-1F category (three R words, one F word), and one 4R-0F category (four R words, no F words). Thus, the categories in a list were split in all possible ways between R words and F words. Every one-fourth of a list contained three R words and three F words.

Across Ss, counterbalancing procedures insured that every word was an R word and an F word equally often and that every category appeared in every category-split condition.

**Procedure.**—The Ss were informed that the lists contained categories; otherwise the procedure was identical to that in Exp. I. The same payoff system was used, and Ss were again asked to circle any words in their final recall that they judged to be F words.

**Results**

Table 3 exhibits the immediate and final recall probabilities of R words and F words in each of the six lists. Once again, in immediate recall, Ss were quite able to recall R words and were equally adept at not recalling F words. Nearly 70% of all R words were recalled and only about 2.5% of F words were intruded. Compared to Exp. I, there was an improvement in R-word recall owing to the categorized nature of the word lists, but the immediate recall of F words remained miniscule.

During the final recall there was a sizable recovery of F words. When Ss were rewarded for every F word they could recall, almost 13% of all F words were retrieved. It appears that Ss were able to use the associations and retrieval cues inherent in the categorized word lists to aid their recall of F words.

Finally, in Table 3, there are little, if any, effects of list position on immediate recall of either F words or R words. However, the effects of list position on final recall performance deserve comment. There appears to be a rather large recency effect across lists for R words and a somewhat smaller recency effect for F words: the probability of recalling words from the last several lists is greater than from the first several lists. This phenomenon is not of major interest, but the close correspondence between R-word and F-word recall is interesting. As R-word recall probability increases, so does the probability of F-word recall. The dependency of F-word recall
on R-word recall is examined in more detail in several subsequent analyses in this section. All subsequent analyses combine the recall data from the six individual lists.

Figure 2 shows the probability of recall for R words and F words, immediately and finally, as a function of category split. It is clear that the probability of recall of an R word or F word in a given category increases in a roughly linear fashion with the number of R words in the category. Without exception, both R-word and F-word recall probability increases with the number of R words in a category.

In Table 4, the probabilities of recalling zero, one, or two F words from a 2R-2F category during the final recall are shown as a function of the number of R words recalled from the category. The message is clear: F-word recall depends on R-word recall. If no R words are recalled, F-word recall is almost zero. On the other hand, if two R words from a single category are recalled, at least one of the F words is recalled in 53% of the cases. Although not shown here, this analysis has also been carried out for the other category splits.

The results of these analyses mirror the results of the 2R-2F analysis reported here.

"Forget" and "remember" labeling.—As in the first experiment, after Ss had completed the final recall task they were asked to circle any item in their final recall that they judged to be an F item. The 48 Ss recalled an average of 42.2 words on the final recall test. Of the 42.2 words, 32.2 were R words, 9.2 were F words, and 0.8 were intrusions. Thus, Ss in this study recalled more R words, more F words, and intruded fewer words from outside the list than did Ss in the first experiment. The Ss circled an average of 9.8 words, a figure once again very close to the actual number of F words recalled. The Ss mistakenly circled only about 8% of the R words and correctly circled 74% of the F words. Slightly more than 50% of the intrusions were circled. Overall, Ss were more accurate in labeling items in this experiment than were Ss in the first experiment.

As in Exp. I, the immediate-recall history of a word heavily influenced whether the word was labeled as an F word or not. Words in S’s final recall that also had been recalled during immediate recall tend to be labeled as R words, and words that were not recalled immediately tend to be labeled as F words, regardless of the actual designation of the words in both cases. Thus, mistakes in labeling consist primarily of R words recalled finally but not immediately and judged to be F words, and F words recalled both immediately and finally and judged to be R words.

**DISCUSSION**

Several aspects of the research reported here merit comment. Perhaps the most striking result is the ability of Ss to "forget" individual items when instructed to do so. In both experiments, F words were intruded in immediate recall at a negligible rate (about 2%), and the recall of R words appeared to suffer little or no interference from F words. Furthermore, Ss were able simultaneously to recall R words and to avoid recalling F words, even though R
words and F words were intermixed during presentation. Past experiments in which a forget instruction has been shown to eliminate interference from F items on the recall of R items (for example, Bjork, 1970b; Bruce & Papay, 1970) have been characterized by a temporal blocking of to-be-remembered and to-be-forgotten items during presentation. The present research demonstrates that temporal blocking of F items and R items is not a necessary condition in order for a forget cue to eliminate or greatly reduce interference owing to F items.

Even though F words were very seldom intruded in the immediate recall of R words and did not appear to interfere with the recall of R words, it is clear that they are not forgotten in the sense of being absent from memory. Instead, it appears that F words exist at some level in memory but are not retrievable unless S's efforts to retrieve them are facilitated in some manner, as, for example, they were by the semantic categorization of the lists in Exp. II. The results of several other recent experiments also illustrate that F items exist in memory. Using a paired-associate probe procedure, Reitman, Malin, Bjork, and Higman (1971) infrequently probed Ss memory for pairs they had been cued to forget and found that Ss were able to respond correctly at a level approximately 60% of their performance on comparable to-be-remembered pairs, although that was only true if Ss were informed by a prearranged signal that they were being tested on an F item; when they were uninformed, recall performance was essentially zero. In experiments by Block (1971) and Elmes et al. (1970), Ss recognition of F items on a delayed recognition test was not significantly worse than their recognition of R items.

The evidence that F words exist in memory rules out the possibility that a forget instruction either prevents storage, erases items from storage, or leads somehow to a very rapid and complete loss of F items from memory. That is, it appears not to be the case that the lack of interference owing to F items during recall of R items is attributable to the nonexistence of F items in memory at the time of recall. The results are consistent, however, with the notion mentioned above that Ss are able, somehow, to organize R items into a grouping that functionally separates them from F items in memory.

The mechanisms by which Ss might accomplish such a differential grouping are not clear. It is somewhat surprising, for example, that Ss in Exp. II had very little difficulty during immediate recall in separating R words in a particular semantic category from F words in the same category. One might have expected that it would be quite difficult to organize a subset of items, all of which were members of a very salient natural category, into an arbitrary grouping that would functionally separate the subset from other items in the category. And if Ss are so good at accomplishing such groupings to facilitate their immediate recall and nonrecall of R words and F words, respectively, it is also somewhat surprising that they seem quite able to use R words to retrieve F words during the final recall. There is no reason to expect, of course, that the mechanisms by which the human information processor differentiates current, to-be-remembered information from information no longer needed are simple-minded.

It is certainly the case in the experiments reported here that Ss rehearsed R words whenever they could and that they devoted little, if any, of their rehearsal time to F words. The selective rehearsal of R words might be assumed to facilitate later recall of R words by increasing the likelihood that R words are stored in long-term memory. It may be, however, that storage in long-term memory is not heavily dependent on rehearsal, but rather, that Ss' attempts to rehearse, relate, and integrate R words have their principal effect on recall by increasing the likelihood that an R word in long-term memory can be successfully retrieved. If that were the case, the systematic nonrehearsal of F words would result in their being peculiarly nonretrievable and, hence, poorly recalled unless retrieval were facilitated in some manner. One implication of such a view, supported in somewhat different experimental contexts by the Block (1971) and Elmes, et al. (1970) experiments referred to above, is that a delayed recognition test should yield a much smaller difference between performance on R words and F words than should a delayed recall test.

It is an interesting possibility that F items are peculiarly noninterfering because they are peculiarly nonretrievable, and vice versa. In other words, the human information processor can selectively group and rehearse to-be-remembered information and thereby protect it from possible interference from to-be-forgotten information, but only at the cost of losing subsequent access to the "discarded" information.
REFERENCES


BJORK, R. A. Control processes and serial position effects in free recall. Paper presented at the Mathematical Psychology meetings, Miami Beach, September 1970. (a)


(Received December 2, 1970)