Constituent Processes in the Differentiation of Items in Memory

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

Three experiments were designed to clarify the perplexing ability of subjects to discriminate between to-be-remembered (TBR) and to-be-forgotten (TBF) items in memory. After the presentation of each pair of words in a list, subjects were first required to solve one to four arithmetic problems and were then cued whether to remember or forget that pair. When subjects were free to use a remember (R) or forget (F) cue in any way they saw fit, their subsequent ability to differentiate TBR and TBF items was impressive, but when subjects were required to retrieve both TBR and TBF word pairs in response to an R or F cue, respectively, their subsequent ability to differentiate TBR and TBF items deteriorated. The results implicate within-list retrieval of TBR items as a potent tagging or strengthening operation that provides a basis on which those items may later be discriminated from TBF items. The "potency" of such events, in terms of their influence on later recall, depends in a clear-cut way on the "depth" (i.e., delay) of retrieval involved, but there are no such comparable effects on later recognition.

It is critical to the reasonably efficient functioning of human memory that we be able to differentiate to-be-remembered information from to-be-forgotten information. In a variety of experiments on directed forgetting, in which cues denote which items are to be remembered and which items are to be forgotten, subjects have shown a remarkable ability to accomplish such differentiation, often under circumstances where items are cued arbitrarily and there is apparently no free time available to use those cues. A particular case in point is the itemby-item cuing procedure introduced by Woodward and Bjork (1971). In this procedure, items (typically, individual words)

are presented one at a time, and subsequent to each item is presented a brief cue as to whether that item is to be remembered (TBR) or to be forgotten (TBF). In cases where the cuing is both completely haphazard (so that TBR and TBF items are neither blocked in time nor differentiable on any apparent basis such as semantic relatedness) and very rapid, subjects are still able to recall about half of the TBR items while intruding only 2 or 3% of the TBF items.

Bjork (1972) has argued that two interrelated processes can account for the whole range of directed-forgetting phenomena. According to Bjork's theory, subjects are presumed to (a) devote all rehearsal and other mnemonic activities to TBR items and to (b) group TBR items in memory in a way that functionally segregates or differentiates them from TBF items. Thus, TBF items do not interfere in the recall of TBR items because they are somehow segregated in memory; they may be recognized as having been presented, but they do not provide retrieval competition.

Although the selective rehearsal and selec-

Copyright 1978 by the American Psychological Association, Inc. 0096-1515/78/0404-0347\$00.75

This report was completed while the first author was a guest of the Human Information-Processing Research Department, Bell Laboratories, Murray Hill, New Jersey. The research reported herein was supported in part by a grant to the first author from the Research Committee of the University of California, Los Angeles, and in part by National Institutes of Health Grant 1F32MH05783-01 to the second author.

Requests for reprints should be sent to Robert A. Bjork, Department of Psychology, University of California, Los Angeles, California 90024.

tive grouping mechanisms together give a reasonable qualitative account of directed forgetting phenomena, there are some situations in which one is hard put to see how subjects could accomplish either selective rehearsal or selective grouping. Consider two examples. In Woodward, Bjork, and Jongeward's (1973) Experiment 3, subjects had to cope with 36-word lists presented 1 word at a time. After each 2.3-sec presentation of a given word, there was a free rehearsal period that ranged from 0 to 12 sec, which was in turn followed by a 1.5-sec within-list test during which subjects were required to recall the current word aloud. Only then were subjects given a 1-sec cue to remember or to forget the current word. The cue was then followed by the next word in the list, and the whole process started all over again. Thus, in the most extreme case, subjects were required to process an item for 15.8 sec without knowing whether they would have to recall the item. Only at the end of that undifferentiated processing were they given a 1-sec cue that they could use for differentiated processing. Whatever the rehearsal period, however, subjects recalled about 40% of the TBR words and intruded about 3% of the TBF words.

Woodward et al.'s results are possibly not so surprising if one assumes that during a given rehearsal period, subjects were able to time-share their rehearsal between the current word and one or two prior TBR words. There was some evidence for such timesharing in that subjects did not always recall the current word on the within-list test. There was also evidence of an indirect sort against such time-sharing in that if Word nin a list was a TBR word, its recall bore no relation to the length of the rehearsal period following Word n + 1. In any case, it is not easy to imagine how subjects were able to achieve the selective processing necessary to recall TBR words quite well while avoiding the intrusion of TBF words.

The problem is more severe in an experiment by Jongeward, Woodward, and Bjork (1975). In Jongeward et al.'s experiment, subjects were presented lists of 32 words. each of which consisted of eight 4-word sets. Each set of 4 words was presented 1 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 to subjects to remember all, only 2 (the first 2 or the last 2), or none of the words in the set. Immediately after the cue, the next word set was presented. Subjects were instructed to devote 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 TBR and TBF 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 4 words they would need to remember. The cue itself was a fairly complex visual array, the encoding of which took essentially all of the 1 sec provided. Subjects were asked not to rehearse words from prior sets during the presentation of later sets, and they said that they did not. Nevertheless, subjects recalled about 35% of the TBR words and they intruded less than 5% of the TBF words.

In our opinion, these results and related results from some different experimental paradigms suggest that there is a missing mechanism in the account of directed forgetting. There are at least three possibilities. First of all, it may be that subjects are simply able to "tag" TBR items in memory quickly and efficiently. Such a tagging notion is unappealing to us because it seems like a somewhat "magical" or tautological solution to the problem of how TBR items are differentiated from TBF items; that is, it seems to assume what one wants to explain. Other researchers (e.g., Anderson & Bower, 1973) have, however, relied heavily on such tagging notions.

A second possibility is that subjects are able to actively inhibit or suppress TBF items. We are not suggesting that subjects might be able to erase TBF items from memory; there is strong evidence against the erasure idea (e.g., Block, 1971; Reitman, Malin, Bjork, & Higman, 1973). It may be, however, that a cue to forget can inhibit or stop a process that would have continued automatically without such a cue. In particular, assume that successful retrieval from long-term memory requires that the memory trace of an item be consolidated in memory. Such consolidation processes would presumably take place without active effort during the period following an item's presentation. It is possible that cues to forget can actively inhibit such retrieval consolidation. The same general idea can be expressed in terms of the levels-of-processing framework (Craik & Lockhart, 1972). If a certain depth of processing is a necessary (but not sufficient) condition for later retrieval from long-term memory, then one might assume that under normal circumstances the required depth is achieved in part via automatic processes that take place subsequent to an item's presentation and initial encoding. In the item-by-item cuing paradigm, one might assume that items are kept at a shallow level of processing through maintenance or primary rehearsal (e.g., Craik & Watkins, 1973; Woodward et al., 1973) until the cue is presented. When a forget cue is presented, it terminates or inhibits the processing that would go on automatically without such a cue.

The third possibility is that the retrieval of an item in reaction to a remember cue is a particularly potent event in terms of strengthening that item's representation in memory and distinguishing it from TBF items. In some sense this idea is the converse of the TBF-inhibition idea. Items are maintained via primary rehearsal at a shallow level of processing until a cue is presented. Items that are designated as TBR items are retrieved and, time permitting, rehearsed in an elaborative or secondary fashion. TBF items are not retrieved, receive no additional processing, and become rapidly inaccessible.

In the first of the experiments reported below, primary rehearsal was nullified by having subjects perform arithmetic problems between the presentation of an item and the R or F cue. In the second experiment, subjects were required to retrieve both TBR and TBF words at the time the cue was presented. The third experiment was designed to clarify the influence of an initial end-of-list test on final recall and recognition.

Experiment 1

The design of Experiment 1 was similar to the design of Woodward et al.'s (1973) Experiment 3, except for the following important changes: (a) Pairs of words rather than single words were presented. This change was intended to decrease the tendency for subjects to rehearse prior TBR items along with the current item. (A pair of words constitutes a much more substantial memory load than does a single word, and mnemonic activities such as imagery and semantic elaboration, which might otherwise be used to interassociate single words presented at different input positions in a list, will tend to be used to interassociate the two words in a given pair instead.) (b) After a word double was presented, subjects were required to solve one, two, three, or four arithmetic problems before a forget or remember cue was presented. Thus, in contrast to Woodward et al.'s experiment, there was no free rehearsal time. (c) As in Woodward et al.'s experiment, there were tests of final recall and recognition for both TBR and TBF items at the end of the experiment, but to clarify the influence of the initial endof-list recall on later recall and recognition, one of the two lists presented was not followed by a test of initial recall.

The design of Experiment 1 was motivated by two considerations. First, as indicated above, we wanted to reduce or eliminate the additional rehearsal TBR items might have received in Woodward et al.'s (1973) experiment owing to the possible time-sharing of rehearsal. Given that such rehearsal is eliminated, it is possible that subjects will no longer be able to differentiate between TBR and TBF items in memory, although the results of the Jongeward et al. (1975) experiment discussed in the introduction would lead one to expect



Figure 1. Initial delayed-recall probability as a function of filled cue delay for the to-be-remembered (TBR) and to-be-forgotten (TBF) items.

otherwise. Second, we wanted to test the idea that the ability of subjects in the Woodward et al. experiment to discriminate between TBF and TBR items in memory depended in part on their having been able to maintain a given item at a superficial level of processing until they were cued to remember or to forget that item. In the present experiment, subjects were not free to simply maintain the current item in short-term memory until the cue appeared. Furthermore, a number of recent studies (e.g., Modigliani, 1976; Whitten & Bjork, 1977) suggest that the more an item's initial retrieval is delayed, the more productive is that act of retrieval in facilitating later retrieval (for a review of retrieval as a learning event, see Bjork, 1975). Thus, whereas Woodward et al. found that recall of both TBR and TBF items was independent of rehearsal time, one would not necessarily expect recall to be independent of the filled cue delay in the present experiment.

Method

Subjects. The subjects were 48 undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles. They were given course credit for their participation.

Materials and procedure. The subjects were tested in groups of three, and all subjects were shown two lists, each beginning with a slide with the word ready on it followed by a series of 16 slides of word pairs. The items were constructed such that all words were common four-letter nonhomophonic nouns, and no pairs contained words that began with the same letter, that rhymed, or that were interassociated in some obvious way. Each word pair was shown for 3 sec. Immediately following each word pair, the subjects were shown either one, two, three, or four simple addition problems, one problem per slide, one slide every 3 sec. All problems were of the form a + b + c =-, where a, b, and c were single-digit numbers. The subjects' task was to solve each problem and write the answers on a sheet of paper. They were told that their performance on the arithmetic task would be scored. After the arithmetic problems, a slide was presented with an F or an R on it to inform the subjects whether to forget or to remember the word pair presented prior to the problems. The subjects were told that there would be a penalty for writing down TBF words on the recall test, that is, their actual recall scores would be equal to the number of TBR words that they wrote down minus the number of TBF words that they intruded. They were also told that they would not be asked to write down any of the words after presentation of some of the lists; that is, they would not be tested on some of the lists.

Prior to beginning the experiment, all subjects were given one list of practice items that was followed by an immediate recall test for the TBR items. The only difference between the practice list and the other lists was that nonsense syllables rather than words were used. This was done to minimize interference between the practice list and the experimental lists of interest.

Immediately after each list had been presented, all subjects were given a series of deductive-reasoning problems to solve for 1 min and 30 sec. Then, half the subjects were asked to write down as many of the TBR words from the preceding list as they could remember, in any order that they wished, while the remaining subjects were asked to solve additional deductive-reasoning problems. One min and 30 sec was allowed for either of these tasks. Each group of subjects was given only one initial delayed-recall test, either after the first list or after the second list.

At the end of the experiment, subjects were asked to free recall as many words from both lists as they could remember, independently of the initial cuing of those words. The test of final free recall was delayed, following the recall or problem solving subsequent to the second list, by an additional 1 min and 30 sec period of deductivereasoning problems. Subjects were allowed 3 min to complete their final free recall. After completion of the test of final free recall, subjects were given a test of final recognition, which consisted of a sheet of paper with 128 words on it-the 64 words used in the experiment plus 64 new words. The subjects' task was to indicate whether each word had been presented in the experiment. There was no time limit on the test of final recognition.

After they had completed the tests of final recall and recognition, subjects were asked to write a short description of (a) what they did during the 3 sec when an R slide was presented, (b) what they did during the 3 sec when an F slide was presented, and (c) whether they tried to remember the word pairs while they solved the arithmetic problems.

Design. After each word pair, and before its R or F cue, there were either one, two, three, or four arithmetic problems corresponding to a filled cue-delay interval of 3, 6, 9, or 12 sec. In each half of the list, eight word pairs were followed by an R cue and eight word pairs were followed by an F cue. Counterbalancing procedures insured that across subjects, each word pair appeared as a TBR or TBF pair equally often, every pair was followed by every filled cue delay equally often, and each list was followed by an initial test or no initial test equally often. The target set (64 words) and the distractor set (64 words) on the test of final recognition were not counterbalanced. The data matrix for the final recall and final recognition measures constituted a $2 \times 2 \times 4$ within-subjects array. The specific factors were presence or absence of an initial test, type of within-list cue (TBR or TBF), and duration of cue delay (3, 6, 9, or 12 sec). The data matrix for the initial delayed-recall measure constituted a 2×4 within-subjects array with the initial-test factor eliminated.

Results

Initial delayed recall. The initial delayedrecall data are shown in Figure 1 as a function of cue (TBR or TBF) and duration of the filled cue-delay interval. An analysis of variance indicated that the main effect of cue was significant, F(1, 47) = 114.4, MS_e = .05, p < .001, with the TBR words being remembered more frequently than the TBF words. The Cue Type × Cue Delay interaction effect also reached significance, F(3, $141) = 2.97, MS_e = .05, p < .05;$ the TBRitem curve in Figure 1 appears to be increasing as a function of cue delay, whereas the TBF-item curve, though subject to possible floor effects, appears to be essentially flat.

Final free recall. The final free-recall data are shown in the bottom panel of Figure 2 as a function of cue, duration of the filled cue-delay interval, and initial testing (initial delayed-recall test on the list or no initial test on the list). An analysis of variance indicated that the main effect of cue was again significant, F(1, 47) = 65.8, $MS_e = .06$, p < .001, with the TBR words being



Figure 2. Final recognition and final free-recall probability as a function of filled cue delay for the to-be-remembered (TBR) and to-be-forgotten (TBF) items from lists that were tested initially (T) and were not tested initially (NT). (The false-alarm probability for new words on the final recognition task was .24.)

remembered more frequently than the TBF words. The main effect of cue delay was also significant, F(3, 141) = 5.8, $MS_e = .04$, p < .001, as was the Cue Type × Cue Delay interaction, F(3, 141) = 3.7, $MS_e = .04$, p < .025. Inspection of Figure 2 suggests that the TBR-item curves are clearly increasing as a function of filled cue delay, whereas the TBF-item curves are essentially flat. These results parallel the initial delayed-recall results.

The Initial Testing × Cue Type interaction did not reach significance, F(1, 47) = 2.7, $MS_e = .04$, p < .10, although it appears in Figure 2 that initial testing facilitated the final recall of the TBR words more than it facilitated the final recall of the TBF words. This result is expected, since more TBR words than TBF words were retrieved on the initial test. Bjork and Woodward (1973) found similar results.

Final recognition. The final recognition data are shown in the top panel of Figure 2 as a function of cue type, filled cue-delay interval, and initial testing. The false-alarm rate was .24. The only significant effect was the main effect of cue type, F(1, 47) = 48.7, $MS_e = .06$, p < .001, with the TBR words being recognized more frequently than the TBF words. All other analyses yielded Fratios less than one. Since Woodward et al. (1973) and Bjork and Jongeward (Note 1) found these two curves to be increasing functions of unfilled primary- or secondaryrehearsal intervals, it does not appear that the present subjects were able to rehearse the words to any appreciable extent during the distractor task. It also does not appear that initial delayed recall of a list facilitated final recognition of the words in that list. Darley and Murdock (1971) also found that final recognition of the items in a list was not influenced by an earlier attempt to recall the items in that list.

Questionnaire data. In response to the question, "What did you do during the 3 sec when an R slide was presented ?" 26 subjects stated that they tried to retrieve the pair of words that immediately preceded the arithmetic problems, while the remaining 22 subjects said that they also tried to associate these two words with previous TBR words. In response to the question, "What did you do during the 3 sec when an F slide was presented?" 22 subjects indicated some sort of active forgetting such as, "I tried to block the words from my mind," 14 subjects said that they did nothing, and the remaining 12 subjects said that they tried to review the TBR words. In response to the question, "Did you try to remember the words while you were required to solve the arithmetic problems?" 34 subjects said that they did not try, while the remaining 14 subjects said that they sometimes tried to remember the words but were unable to do so because of the time constraints.

Discussion

The results of Experiment 1 have some implications as to the possible "missing mechanisms" in directed forgetting outlined in the introduction; but before those implications are discussed, there are several other aspects of the results that merit comment. First, it is interesting to note that, taken together, the effects of the unfilled cue delay in Woodward et al.'s (1973) Experiment 3 and the effects of the filled cue delay in the present experiment do serious violence to commonsense predictions. Woodward et al. found that recall of TBR words did not increase as amount of free rehearsal varied from 0 to 12 sec. In the present experiment, however, recall of TBR words did increase as a function of the amount of interfering activity that intervened between the presentation of a word pair and the cue to remember that pair. Apparently, the more an act of retrieval is delayed, the greater the depth of processing that is required and the more such a retrieval effort will facilitate later efforts to retrieve. This result is, of course, consistent with the recent findings of other researchers (e.g., Jacoby, 1973; Modigliani, 1976; Whitten & Bjork, 1977). Retrieval depth, on the other hand, did not appear to influence the final recognition of TBR words, nor did the presence or absence of an initial end-of-list free recall influence later recognition of TBR words. Retrieval profits substantially from prior efforts to retrieve; recognition does not.

As far as the differentiation of TBR and TBF items is concerned, the present results are clearly consistent with the notion that TBR items are retrieved in reaction to a remember cue and that the act of retrieval is a potent event that increases the strength of TBR items in memory relative to TBF items (which are not retrieved in reaction to a forget cue). The results do not have much to say with respect to the tagging idea. It is possible, of course, that the act of retrieving TBR items in response to a remember cue does not so much increase the strength of TBR items relative to TBF items as tag those items as TBR items. If, however, final recognition is taken to be a reasonable measure of an item's strength in memory, then the TBR words in Experiment 1 were clearly stronger than the TBF words.

Finally, the results provide no support for the TBF-item inhibition or suppression idea. One might have expected that a cue to forget would be increasingly less effective from an inhibitory standpoint the more it was delayed following the presentation of an item. If there is an automatic consolidation-type process that takes place subsequent to an item's presentation, and if a forget cue can inhibit or suppress any further such consolidation, then TBF-word recall should have increased with the number of intervening arithmetic problems. In fact, none of the measures of a subject's memory for TBF items-intrusions in initial free recall, final free recall, and final recognition-varied as a function of the cue delay in the present experiment. It appears that the only processing devoted to TBF items took place during the initial 3-sec presentation. Typically, that processing would be relatively superficial. (Given that one may later be cued to forget the current word double, it is counterproductive to work hard at achieving a strong encoding of those words in memory until one is cued to remember them.) In cases where two words had some idiosyncratic association for a given subject, of course, some relatively automatic semantic processing no doubt took place; such semantic processing was probably the principal basis for the successful final free recall of TBF words. From subjects' reports, they were either unable or did not try to rehearse items during the arithmetic problems; and in reaction to a forget cue, they did nothing with respect to the current word double, although they sometimes tried to retrieve and rehearse prior TBR words.

Experiment 2

If there is any validity in the foregoing account of the results of Experiment 1, then requiring subjects to retrieve the current word double in response to either a forget or remember cue should substantially reduce the discriminability of TBR and TBF items in memory. Experiment 2 was designed to explore that implication.

Method

The design of Experiment 2 was identical to the design of Experiment 1 except for one important procedural innovation. The 48 subjects in Experiment 2 were asked to retrieve the current word double in response to both forget cues and remember cues. They were given a sheet of paper that had a left-right pair of small boxes corresponding to each word pair in a list. When an R or F cue was presented, subjects were asked to check the left-hand or right-hand box if they could still remember the left-hand or right-hand member, respectively, of the current word pair. They had to complete this task within the 3 sec during which an R or F slide was shown. They were asked to indicate honestly what words they could and could not remember, but they were told, as in Experiment 1, that their recall score at the end of a list would be the number of TBR words they recalled minus the number of TBF words they intruded.

The box-checking procedure was used rather than having subjects write down the words they could remember because written responses would have necessitated an increase in the length of time a cue was presented. Increasing the cue duration would have increased the time available for selective rehearsal of TBR words and, in general, would have decreased the comparability of Experiments 1 and 2. From the subjects' standpoint, the box-checking procedure seemed quite natural and was not difficult to execute in a reliable fashion.

Results

Within-list recall. In the top panel of Figure 3 is plotted the proportion of cases in which subjects indicated (by box checking) that they could still recall a given word at the end of the filled cue-delay interval. Performance is shown as a function of cue type (TBR or TBF), cue delay, and whether there was or was not an initial delayed test. As one would hope, the only significant effect on performance was that of cue delay, F(3, 141) = 51.5, $MS_e = .05$, p < .001. If subjects were following instructions, the nature of the cue should not have mattered, nor, of course, should the presence or absence of an end-of-list test.

Initial delayed recall. Performance on the test of initial delayed recall is shown in the bottom panel of Figure 3 as a function of cue type and cue delay. The influence of cue type was significant, F(1, 47) = 5.3, $MS_e = .03$, p < .05, as was the influence of



Figure 3. Top panel: Within-list recall probability as a function of filled cue delay for the to-beremembered (TBR) and to-be-forgotten (TBF) items from lists that were tested initially (T) and were not tested initially (NT). Bottom panel: Initial delayed-recall probability as a function of filled cue delay for the TBR and TBF items.

cue delay, F(3, 141) = 9.4, $MS_e = .05$, p < .05.001. The Cue Type \times Cue Delay interaction was not significant (F < 1). Because of a possible floor effect, one does not know whether the influence of cue delay is greater than or lesser than the apparent increase in Figure 3, which seems limited to an advantage of the 12-sec delay over the 3-, 6-, and 9-sec delays. In any case, there is a striking contrast in the top and bottom panels of Figure 3. Although the frequency of within-list recall decreases with increasing cue delay, the increased potency of successful retrievals at the longer delays more than offsets that disadvantage in terms of later recall.

There is also a striking contrast in the levels of initial delayed recall in Experiments 1 and 2. Having to retrieve TBF words as well as TBR words in response to the corresponding cue dramatically impaired the ability of subjects in Experiment 2 to differentiate TBR and TBF items in memory. Compared to the subjects in Experiment 1, the subjects in Experiment 2 recalled only about half as many TBR words, and they intruded about twice as many TBF words. Also, the recall of TBF words, which was independent of cue delay in Experiment 1, increased with cue delay in Experiment 2.

Final free recall. The final free-recall data are shown in the bottom panel of Figure 4 as a function of cue type, cue delay, and whether there was an initial delayed test of the list. The analysis of variance indicated that the main effect of cue type was not significant (F < 1). The words that the subjects were instructed to forget were just as accessible in long-term memory as the words that they were told to remember. Apparently, the difference between TBR- and TBF-word recall in the initial delayed recall of the TBR words was due to the suppression of some TBF words. The recall of the TBF words from lists that were tested initially increased slightly from .11 in initial delayed recall to .16 in final free recall, whereas the recall of the TBR words stayed about the same, .16.

The main effect of initial testing of a list was significant, F(1, 47) = 12.4, $MS_e = .05$, p < .001, with the words being more likely to be recalled in final free recall if the list was tested initially. That this was the case to the same extent for the TBR and TBF words indicates that the words did not have to be written down in the initial test to receive retrieval facilitation for the later test. Items that were initially retrieved but then suppressed also received the benefit of the initial testing.

The main effect of cue delay was also significant, F(3, 141) = 13.6, $MS_e = .04$, p < .001, with the probability of final free recall of both the TBR and TBF words being greater the longer the filled cue delay. In Experiment 1, only the TBR words showed an increase in final free recall as a function of cue delay. Hence, initial retrieval depth influences later retrieval whether subjects have any intent to retrieve the item or not.

Final recognition. The final recognition data are shown in the top panel of Figure 4 as a function of cue type, cue delay, and whether a list was initially tested. The false-alarm rate (.21) was comparable to that obtained in Experiment 1. Unlike Experiment 1, there was no difference between the final recognition of the TBR and TBF words. An analysis of variance indicated that only the interaction between cue delay and whether a list was tested initially was significant, $F(3, 141) = 2.7, MS_e = .04, p$ < .05. Inspection of Figure 4 suggests that this interaction arises because of an unaccountable aperiodicity in the test (T) and no-test (NT) curves.

Questionnaire data. The 48 subjects in this experiment were asked, "What, if anything, did you do differently when an R slide was shown than when an F slide was shown?" Thirty subjects stated that they did not do anything differently, whereas the remaining 18 subjects stated that they sometimes tried to remember previous TBR words in addition to the pair of words just presented but were usually not able to do so because of the time constraints. In response to the question, "Did you try to remember the words while you were required to solve the arithmetic problems?" 36 subjects said that they did not, and the remaining 12 subjects said that they sometimes tried but were usually not able to do so. The subjects' responses to the latter question were similar to those obtained from the subjects in Experiment 1.

Discussion

The results of Experiments 1 and 2 provide strong support for the notion that differences in the frequency of within-list retrievals of TBR and TBF items are an important basis on which those items are later differentiated in memory. A within-list retrieval strengthens an item's representation in memory, with the extent of such strengthening clearly dependent on the de-



Figure 4. Final recognition and final free-recall probability as a function of filled cue delay for the to-be-rememberd (TBR) and to-be-forgotten (TBF) items from lists that were tested initially (T) and were not tested initially (NT). (The false-alarm probability for new words on the final recognition task was .21.)

lay of that retrieval from the item's initial presentation. When subjects were forced to retrieve both TBF and TBR items in Expriment 2, there was no ultimate difference in the strengths of those items in memory as indexed by final free recall and recognition. The results of Jongeward et al. (1975) discussed at the beginning of this article, which apparently demonstrated a mystifying ability on the part of subjects to discriminate TBR and TBF items, no longer seem so mysterious if one assumes that the subjects in their experiment retrieved only the TBR items in response to the within-list cues.

It is interesting to contrast the results of Experiment 2 with the results of Woodward et al.'s (1973) Experiment 3. In both experiments, a within-list retrieval of both



Figure 5. Top panel: Within-list recall probability as a function of filled cue delay for the to-beremembered (TBR) and to-be-forgotten (TBF) items from lists that were tested initially (T) and were not tested initially (NT). Bottom panel: Initial delayed-recall probability as a function of filled cue delay for the TBR and TBF items.

TBR and TBF items was required. Why, then, were subjects in the Woodward et al. experiment able to discriminate so well between TBR and TBF items in memory? In our view, there are three reasons. First of all, subjects in Woodward et al.'s experiment were able to devote additional rehearsal to TBR items by time-sharing their free rehearsal time between the current item and TBR items presented earlier. Second, the within-list retrievals in Woodward et al.'s experiment were not the potent events that the within-list retrievals were in the present

Experiments 1 and 2. In Woodward et al.'s Experiment 3, the within-list retrieval of a given item followed an unfilled interval during which subjects were keeping the item available via primary rehearsal. The act of retrieval was, therefore, typically a trivial event, and one might expect that its longterm consequences would be quite insubstantial. Finally, the cue to remember or forget was presented after the within-list retrieval had taken place. Thus, during the presentation of either cue, subjects could use the time available to increase their processing of TBR items. No such additional time was available in the present Experiment 2.

It is also interesting to note in Experiment 2 that the subjects' intent with respect to the items retrieved on a within-list test had little to do with the consequences of that retrieval. In the same vein, Bugelski (1970) has shown that requiring complex processing of TBF words, such as imaging, eliminates the deleterious effect of forget cues on the recall of TBF items. One cannot, however, dismiss the notion that subjects can tag items in memory as TBR or TBF items in some qualitative fashion. The significant difference in the initial delayed recall of TBR and TBF items in Experiment 2 provides compelling evidence for such a tagging process. Given that TBR and TBF items did not differ in strength, as is evidenced by their equivalent levels of final recall and recognition, subjects must have been able to suppress the initial delayed recall of TBF items based on a qualitative tag of some kind. The tagging notion also seems more plausible in view of results like those obtained by MacLeod (1975), who found that subjects could still discriminate TBR and TBF items on a recognition test administered 2 weeks after the list had been presented.

Experiment 3

Experiment 3 was designed both to replicate Experiment 2 and to clarify the influence of initial delayed recall on final free recall. In both Experiments 1 and 2, whether the first list was tested initially influenced final free recall of that list substantially, but whether the second list was tested initially had little or no effect on the final free recall of that list. In Experiment 2, there was another unexpected interaction with list. There was a much larger difference between TBR and TBF words in the levels of their initial recall following List 1 than there was following List 2. Neither of these list effects appeared to interact with cue delay.

It is not particularly useful to examine or discuss in detail the unexpected list effects obtained in Experiment 1 and 2. The designs of Experiment 1 and 2 make it impossible to determine the source of such list effects. If the first list was tested, the second was not, and vice versa; the authors did not foresee that which list was tested initially would interact with any other variables. In Experiment 3, whether the first list was tested initially was crossed with whether the second list was tested initially.

Method

Subjects. The subjects were 48 undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles.

Procedure. The procedure was the same as that used in Experiment 2, except for three changes. (a) Rather than having one, two, three, or four arithmetic problems after the presentation of a given word double, there were either two or four problems. (b) To avoid the possible floor effects obtained in Experiment 2, the lists were shortened from 16 to 12 word pairs. (c) There were 12 subjects in each of four initial-test conditions: T-T, T-NT, NT-T, and NT-NT, where the first and second symbols denote whether there was an initial test (T) or no initial test (NT) after the first and second lists, respectively.

Design. Counterbalancing procedures were instituted as before to insure that each word pair appeared as a TBR or TBF pair equally often, and every pair was followed by each filled cue delay equally often. The data matrix for the withinlist recall, final free recall, and final recognition measures was a $2 \times 2 \times 2 \times 2 \times 2$ mixed design. The specific factors were initial testing of the first list (initial delayed-recall test or no initial test), initial testing of the second list, list (first list or second list), cue type (TBR or TBF), and filled cue-delay interval (6 or 12 sec). Initial testing of the first and second lists were the only betweensubjects factors.

The initial delayed-recall data, obtained from



Figure 6. Final recognition and final free-recall probability as a function of filled cue delay for the to-be-remembered (TBR) and to-be-forgotten (TBF) items from lists that were tested initially (T) and were not tested initially (NT). (The false alarm probability for new words on the final recognition test was .21.)

subjects who received at least one initial test, were analyzed separately for the first list and the second list. For each of these two analyses, the data matrix was a $2 \times 2 \times 2$ mixed design. The specific factors were initial testing (initial delayed-recall tests given on only one list or on both lists), cue type (TBR or TBF), and filled cue delay (6 or 12 sec). The only between-subjects factor was initial testing.

Results

Within-list recall. The probability of within-list recall, as determined by the number of boxes checked by the subjects, is shown in the top panel of Figure 5 as a function of cue type, cue delay, and whether a list was tested initially. The subjects checked slightly more boxes during the second list than during the first list (82% vs. 79%) but the difference was not significant. As in Experiment 2, the only significant effect was the main effect of cue delay, F(1, 44) = 73.4, $MS_e = .06$, p < .001; the probability of remembering the words was smaller with the longer cue delay, but the subjects were just as likely to retrieve the words during the presentation of an F slide as during the presentation of an R slide.

Initial delayed recall. Since the pattern of results was the same for the first and second lists, the initial delayed-recall data from both lists were combined in the bottom panel of Figure 5 as a function of cue type and cue delay. The probability of recalling the TBR items was .31 on the first list and .35 on the second list, whereas the probability of recalling the TBF items was .15 on the first list and .14 on the second list. The main effect of cue type was significant for both the first list, F(1, 22) = 18.8, $MS_e =$.03, p < .001, and the second list, F(1, 22) $= 55.2, MS_{e} = .02, p < .001$. The main effect of cue delay was also significant for both the first list, F(1, 22) = 29.3, $MS_e =$.01, p < .001, and the second list, F(1, 22) $= 4.3, MS_{e} = .02, p < .05.$ Finally, initial delayed recall did not decrease from the first list to the second list (.23 vs. .25), and initial delayed recall on the second list was not affected by whether the first list was tested initially (.25 vs. .24). Therefore, it does not appear that proactive interference was a factor in this experiment in the initial delayed recall. In fact, there were no words from the first list given as intrusions during the initial delayed-recall tests on the second list. This lack of intrusions was no doubt in part (or completely) attributable to there being 3 min of activity between the end of the first list and the beginning of the second list.

Final free recall. The final free-recall data are shown in the bottom panel of Figure 6 as a function of cue type (TBR or TBF) and cue delay for lists that were either tested initially or were not tested initially. The main effect of cue delay was

significant, F(1, 44) = 106.8, $MS_e = .02$, p < .001, with the probability of final free recall of both the TBR and TBF words being greater with the longer cue delay. In addition, the Initial Testing on the First List \times List interaction was significant, F(1, $(44) = 31.0, MS_e = .01, p < .001$, as was the Initial Testing on the Second List \times List interaction, F(1, 44) = 22.5, $MS_e = .01$, p < .001. Cicchetti (1972) tests conducted on these interactions indicated that more words were recalled from the first list if it was tested initially (p < .01), and more words were recalled from the second list if it was tested initially (p < .01). Thus, final recall was greater for lists that were tested initially.

Unlike initial delayed recall, the main effect of cue (TBR or TBF) was not significant in final free recall. The difference between TBR- and TBF-word recall during the initial tests was in all likelihood due to suppression of the TBF words, as was concluded in Experiment 2. The probability of TBF-word recall increased from .14 in initial delayed recall to .24 in final free recall, whereas the probability of TBR-word recall decreased from .33 in initial delayed recall to .26 in final free recall.

Final recognition. The final recognition data are shown in the top panel of Figure 6 as a function of cue type and cue delay for lists that were tested initially and were not tested initially. The false-alarm rate was identical to that obtained in Experiment 2, .21. The analysis of variance showed no significant sources of variation.

Influence of initial-test condition on final free recall. In Figure 7, the final free-recall data are shown as a function of the initialtest conditions. The data in Figure 7 are quite striking. They demonstrate the operation of three remarkably stable factors in determining the level of final free recall, none of which interacts with the others. First of all, the advantage of the top two curves over the bottom two curves shows the advantage of list recency. The analysis of variance indicated that the main effect of list was significant, F(1, 44) = 192.7, MS_e = .01, p < .001, with more words recalled from the second list than from the first list. Second, the advantage of the right-hand points over the left-hand points shows the advantage of retrieval practice, an effect that has already been mentioned above. Finally, there is an advantage in the final recall of a given list if the other list was not tested initially (ps < .05), an effect that is shown by the differences between the two curves for a given list. It appears that the initial recall of a given list strengthens that list in memory with respect to final recall, not only improving the final recall of that list but also impairing the recall of the other list via increased retrieval competition or some similar process.

For each of the three effects in Figure 7 recency, retrieval practice, and retrieval competition—the data provide four independent estimates of the size of that effect. In each case it is remarkable how consistent those four estimates seem to be, as can be seen from Figure 7.

Discussion

Experiment 3 was quite productive. On the one hand, the major results of Experiment 2 were clearly replicated. Within-list recall decreased as a function of cue delay, whereas initial delayed recall and final free recall increased as a function of cue delay for both TBR and TBF items. Also, there was once again evidence that the TBF and TBR items were differentiated in memory on some qualitative basis that allowed subjects to suppress some of the TBF items during initial delayed recall. The strength of TBR and TBF items in memory, as judged by final free recall and recognition, did not differ. Finally, as in Experiment 2, final recognition was unaffected by cue type, cue delay, and whether a list was initially tested.

On the other hand, some possibly problematic list effects that appeared with the incompletely crossed design of Experiment 2 were not replicated. Rather, Experiment 3 provided an unusually clear picture of the benefits attributable to retrieval practice and the costs attributable to retrieval competition, as shown in Figure 7.



Figure 7. Final free-recall probabilities as a function of initial-test conditions. To facilitate the interpretation of these results, the influence of recency (A), retrieval competition (B), and retrieval pracice (C) is indicated.

General Discussion

The experiments reported herein have implications with respect to two important and apparently somewhat interrelated problems; (a) how TBR and TBF items are differentiated in memory and (b) how initial retrieval events influence later recall and recognition. The two problems are apparently interrelated because the present results provide strong evidence that the eventual discriminability in memory of TBR and TBF items is largely attributable to the initial retrieval of TBR items and nonretrieval of TBF items in response to R cues and F cues, respectively. When the situation requires subjects to retrieve TBF items as well as TBR items, as in Experiments 2 and 3, their ability to differentiate TBR and TBF items is greatly reduced, and tests in which both TBR and TBF items are to be recalled or recognized reveal no differences in the relative strengths in memory of TBR and TBF items.

As pointed out in the discussion of Experiments 2 and 3, however, there is more to the differentiation of TBR and TBF items in the item-by-item cuing situation than differential within-list retrieval of TBR items. Subjects are able to suppress TBF items in initial delayed recall even though those items have apparently the same strength in memory as TBR items. There is, therefore, evidence for a qualitative tagging process of some kind. Such a tagging process could take the form of storing TBR and TBF items in somewhat different areas in memory-a notion we find more appealing than the tagging notion, although it may be formally equivalent. If TBR items were interassociated with each other, and TBF items were not, then that would provide a reasonable qualitative basis for differentiation. It seems unlikely, however, that any such differential organization explains the suppression of TBF items in Experiments 2 and 3. First of all, the experimental procedure does not seem to provide the time necessary for subjects to process more than the current word pair. And second, if TBR items had the advantage of being interassociated, then such interassociation should also facilitate the final free recall of the TBR items—a result that was not obtained. For the moment, it appears that we are stuck with the tagging hypothesis. The hypothesis that an F cue suppresses or inhibits a consolidation-type process that would take place without such a cue-a notion toward which we felt predisposed prior to this researchis not supported by any aspect of the results reported herein.

The influence of initial retrieval on later recall and recognition has been discussed in some detail earlier in this article, and those points will not be reiterated here. There are two curious aspects of the results, however, that merit additional comment. One has to do with the influence of initial retrieval on later recognition. In all three experiments, there was no influence on final recognition of either within-list retrieval depth or whether a list had an initial delayed recall test. However, given our interpretation of subjects' activities during the presentation of the R and F cues in Experiment 1, the fact that the final recognition of TBR words in Experiment 1 clearly exceeded the final recognition of TBF words poses something of a puzzle. In our view, the only difference in the processing of TBR and TBF words in Experiment 1 was that TBR words were retrieved in response to an R cue, whereas TBF words were not retrieved in response to an F cue. If that act of retrieval (of the TBR words) facilitated later recognition, why was there no influence of an initial delayed test on later recognition? We interpret this puzzle as follows: The recognition task used in these experiments required a judgment as to whether given words (all of which were common and, hence, familiar) had or had not occurred in the experimental context. We assume that any occurrence of a word in the experiment (whether it is presented, rehearsed, or retrieved on a within-list test) helps to associate that word with the situational context, which in turn facilitates recognition. The reason that an end-of-list delayed test does not help final recognition is that the words retrieved on such a test tend to be a subset of those that would be recognized on an end-of-experiment test anyway. That is, successful retrieval of an item on an end-of-list test requires greater "strength" in long-term memory than does end-of-experiment recognition. Successful retrieval on a within-list test, however, does not require that an item already have enough strength in long-term memory to support final recognition; hence, a within-list retrieval can provide additional, possibly important, occurrence information.

The second aspect of the results that deserves additional comment is the evidence in Experiment 3 that testing a given list not only improves its later recall but also increases the retrieval competition provided by that list in the recall of the other list. That an initial retrieval will facilitate later retrieval is not news (e.g., Bjork, 1975; Bjork & Woodward, 1973; Cooper & Monk, 1976), but that an initial retrieval will provide later retrieval competition in the recall of other items is a new and interesting result. It has been demonstrated before that strengthening some items in a list by additional presentations will decrease the recall of the other list items (Tulving & Hastie, 1972). Whether such effects occur, however, appears in part to be determined by whether the strengthened items are in the same

search set as the items being retrieved. In Experiment 3, initial delayed recall of the second list did not depend on whether the first list had been recalled. Thus, when subjects were trying to retrieve only the items in the second list, there was little or no effect of whether the first-list items had been strengthened by an initial delayed test. However, from the standpoint of the final free-recall test, when both lists fell within a common target set at the time of recall, increasing the accessibility of one list of items (via an initial free-recall test) had negative effects on the retrieval of the items in the other list.

The present research started out as a search for a "missing mechanism" in directed forgetting. The results, as discussed above, turned out to have implications with respect to a variety of other problems as well, such as differences between recognition and recall, the influence of initial retrieval events on later retrieval, and so forth. With respect to the original question, the results point to retrieval events during the presentation of R and F cues as the locus of important qualitative and quantitative processes that underlie the striking ability of subjects to differentiate TBR and TBF items during recall.

Reference Note

1. Bjork, R. A., & Jongeward, R. H., Jr. Rote versus elaborative rehearsal as updating processes. Paper presented at the meeting of the Psychonomic Society, Denver, Colorado, November 1975.

References

- Anderson, J. R., & Bower, G. H. Human associative memory. Washington, D.C.: Winston, 1973.
- Bjork, R. A. Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), Coding processes in human memory. Washington, D.C.: Winston, 1972.
- Bjork, R. A. Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), Information processing and cognition. Hillsdale, N.J. : Erlbaum, 1975.
- Bjork, R. A., & Woodward, A. E. Directed for-

getting of individual words in free recall. Journal of Experimental Psychology, 1973, 99, 22-27.

- 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.
- Cicchetti, D. V. Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. Psychological Bulletin, 1972, 77, 405-408.
- Cooper, A. J. R., & Monk, A. Learning for recall and learning for recognition. In J. Brown (Ed.), Recall and recognition. New York: Wiley, 1976.
- Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.
- Craik, F. I. M., & Watkins, M. J. The role of rehearsal in short-term memory. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 599-607.
- Darley, C. F., & Murdock, B. B., Jr. Effects of prior free-recall testing on final recall and recognition. Journal of Experimental Psychology, 1971, 91, 66-73.
- Jacoby, L. L. Encoding processes, rehearsal, and recall requirements. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 302-310.
- 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.
- MacLeod, C. M. Long-term recognition and recall following directed forgetting. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 271-279. Modigliani, V. Effects on a later recall by delay-
- ing initial recall. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2. 609-622.
- 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.
- Tulving, E., & Hastie, R. Inhibition effects of Intralist repetition in free recall. Journal of Experimental Psychology, 1972, 92, 297-304.
 Whitten, W. B., & Bjork, R. A. Learning from tests: Effects of spacing. Journal of Verbal
- Learning and Verbal Behavior, 1977, 16, 465-478.
- Woodward, A. E., & Bjork, R. A. Forgetting and remembering in free recall: Intentional and unintentional. Journal of Experimental Psychology, 1971, 89, 109-116.
- Woodward, A. E., Bjork, R. A., & Jongeward, R. H., Jr. Recall and recognition as a function of primary rehearsal. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 608-617.

Received December 6, 1977