Journal of Experimental Psychology: Human Learning and Memory
1975, Vol. 104, No. 3, 271–279

Long-Term Recognition and Recall Following Directed Forgetting

Colin M. MacLeod
University of Washington

Two experiments investigated long-term retention subsequent to directed forgetting. In Experiment 1, both recognition and cued recall were better for categories given remember (R) instructions than for categories given forget (F) instructions. A constant advantage of R items over F items persisted from an immediate test to a 1-wk delayed test. In Experiment 2, recognition of R items exceeded that of F items at retention intervals of 1 and 2 wk, the superiority of R items over F items again being constant across retention intervals. Presence or absence of study-instruction cues at the time of test in Experiment 2 did not differentially affect performance. An explanation is offered relating the directed forgetting effect to selective rehearsal during initial processing of the items and to the presence of instruction information stored with the individual items.

The directed forgetting paradigm, introduced by Muther (1965) and usually dated from a study by Bjork, Laberge, and Legrand (1968), has become firmly established as a means of examining "intentional" forgetting, i.e., forgetting that the subject has been told to do (for reviews see Bjork, 1972; Epstein, 1972). Essentially, the directed forgetting paradigm involves instructing the subject to forget certain items (F items) and to remember the remaining items (R items) in a list. The subject is told that he will not subsequently be tested on F items, so that it is to his advantage to devote his processing only to the R items. Numerous studies (e.g., Bjork, 1970; Block, 1971) have shown that performance on R items in such a list is improved relative to performance on the same items in a list in which all items are to be remembered. Furthermore, when F items themselves are unexpectedly tested, they show a clear performance decrement (e.g., Davis & Okada, 1971; Woodward & Bjork, 1971), at least when measured by recall. Given these standard findings, the "directed forgetting effect" will be defined in this article as a greater proportion of correct R items than of correct F items on a retention test.

The question of where the F instruction exerts its influence in the course of processing is central. One interpretation suggests that the directed forgetting effect stems primarily from differences during initial processing (the "selective rehearsal" account). Two types of strategies could be employed in this case: (a) Each item is processed elaboratively from the time it is presented, but elaboration is discontinued for F items when the F instruction is presented (cf. Craik & Watkins, 1973); and (b) items are maintained by rote repetition until their related instruction is presented, at which time only the R items receive further (elaborative) processing (cf. Woodward, Bjork, & Jongeward, 1973).

An alternative conception (the "selective search" account) contends that the directed forgetting effect is due primarily to the discrimination of instruction sets after they have already been stored. Specifically,

A shorter version of Experiment 1 was reported at the meeting of the American Psychological Association, Montreal, August 1973. I thank Bob Weltzien (Experiment 1) and Connie Copeland (Experiment 2) for collecting the data. I am especially grateful for the generous advice and encouragement of my committee chairman, Thomas O. Nelson, throughout this research.

Requests for reprints may be sent to Colin M. MacLeod, who is now at the School of Social Sciences, University of California, Irvine, California 92664.

selective search contends that the R items and F items are separately tagged, and that there is a higher priority of searching among the R items. However, as Bjork (1970, 1972) has pointed out, a comprehensive explanation of the effect may require invoking both accounts, since both rest on differentiating R and F items (although it would appear that the instruction label must be retained only in the selective search account). Determining the extent of involvement of selective rehearsal and of selective search is one central aim of the present research.

Unfortunately, the evidence is equivocal for both selective rehearsal and selective search. In the case of selective rehearsal, additional opportunities to rehearse should allow the subject to accentuate the advantage of R items over F items. Epstein and Wilder (1972) have demonstrated such a directed forgetting effect, but earlier studies by the same researchers (Epstein, Massaro, & Wilder, 1972; Shebilske, Wilder, & Epstein, 1971), as well as by Woodward and Bjork (1971), did not obtain the predicted accentuation. Still, as Epstein and Wilder point out, their evidence is the most direct because they used a within-subjects design, whereas the earlier studies varied rehearsal opportunities between subjects.

The Woodward et al. (1973) study has important implications for theories of directed forgetting, particularly those involving the selective rehearsal notion. Their critical finding is that the type of rehearsal (maintenance or elaboration) is more important than the amount of rehearsal. Whereas longer maintenance rehearsal apparently enhances only recognition performance, increased elaborative rehearsal seems to enhance both recognition and recall performance. This rehearsal distinction will be referred to throughout the present article.

Another set of conflicting findings occurs in the recognition studies focusing on selective search. If the subject has distinct R and F sets in storage, and if the superiority of R items over F items is due to a greater probability of instituting a search in (and thus retrieving items from) the R set, this can account for the directed forgetting effect in recall. However, since retrieval difficulty is at least minimized in recognition (e.g., Kintsch, 1970; Murdock, 1968), the directed forgetting effect should be reduced when measured by recognition. Block (1971), Elmes and Wilkinson (1971), and Elmes, Adams, and Roediger (1970) did find equal recognition of R and F items; however, Davis and Okada (1971) obtained better recognition of R items than of F items. The Davis and Okada study warrants considerable weight in this controversy because it was specifically designed to rectify problems in the earlier recognition studies. Nevertheless, for the directed forgetting effect to be understood, conflicting results such as these must be resolved. The Woodward et al. (1973) findings suggest one step toward resolution of this conflict. Perhaps, in those studies where recognition of F and R items was equivalent, equality of maintenance rehearsal was the determining factor. It may be that R-item accentuation occurs in recognition only to the extent that R items receive considerably more elaborative rehearsal than do F items. This would be possible if maintenance rehearsal affects primarily item storage, whereas elaborative rehearsal affects primarily retrieval, with a lesser effect on storage. Since it is assumed that only R items are elaborated, a considerable amount of elaboration would be required to produce the directed forgetting effect in recognition (where retrieval benefits are minimal). However, the effect would appear in recall even with a slight elaboration advantage for R items because retrieval plays a more integral role in recall.

The present experiments examine long-term retention, measured by both recognition and recall, of different types of material following directed forgetting. Both experiments involve intraserial F instructions during acquisition, followed by delayed tests of F and R items. The problems inherent in testing F items (Bjork, 1972) are eliminated by the use of only one extended test trial so that the credibility of the F instruction during acquisition is not undermined.

Experiment 1

Investigations of the directed forgetting effect have typically been conducted in only one experimental session; consequently, estimation of the stability of the effect over time has not been possible. With a between-subjects design, retention can be measured after various delays without the contaminating effect of repeated tests on the same subject. Using such a design, the present experiment examined both recall and recognition immediately after acquisition (0 wk) or after a delay (1 wk).

The presence of a directed forgetting effect in recall would appear to be a prerequisite for the examination of recognition effects. Then if there is a negligible effect in recognition, the mechanism for discriminating F items from R items in storage should be central in theories of directed forgetting. Of course, differential storage requires that instruction information be stored during original processing; however, the implication is that only retrievability, and not storage, would be affected. If, on the other hand, the directed forgetting effect is present in recognition, then the role of initial processing should be emphasized. An R-item advantage in recognition implies that F items are not stored to the same degree as R items. Thus, the recognition data will permit a comparison of the selective search and selective rehearsal accounts.

Method

Design. The design was a 2 × 2 × 2 factorial with two between-subjects variables and one within-subjects variable. The major between-subjects variable was retention interval (0 or 1 wk). The other between-subjects variable, study list, was included to counterbalance the instruction presented with a given category at study. The within-subjects variable was instruction, either R or F, occurring with each category during list presentation.

Subjects. The subjects were 48 University of Washington undergraduates whose participation partially fulfilled a course requirement. Twelve subjects were randomly assigned to each of the four groups defined by factorial combination of retention interval and study list. The subjects participated individually at both study and test.

Stimuli. The 60 study items and the 96 distractors (for the recognition test) were selected from the ranks 10–32 in the Battig and Montague (1969) category norms. From each of 16 categories, 3 study items and 6 distractors were chosen. From another 4 categories, only the 3 study items were chosen. The three exemplars from each of the 20 studied categories were selected such that they did not begin with the same letter. Similarly, on the recognition test, each set of 2 distractors and 1 study item could not begin with the same letter. Otherwise, assignment of distractors to study items was random within the category.

The four categories without distractors were used as "buffer" categories (two at the beginning and two at the end of the study list). These buffer categories served two purposes: (a) reduction of list-related serial position effects on the 16 critical categories and (b) decoy test items immediately after list presentation.

Procedure. The following blocked sequence was used for list presentation: category name (3 sec), 3 category exemplars (3 sec each), and instruction, either *remember* or *forget* (6 sec). A deck of 5 × 8 in. (13 × 20 cm) index cards was used for list presentation. Category names were presented on blue cards, category exemplars on white cards, and instructions on orange cards. All words were typed in .25-in. (.64-cm) uppercase bulletin type for easy reading.

The 4 buffer categories always had R instructions. Of the 16 remaining critical categories, half had R instructions and half had F instructions. Assignment of instruction was random with the restriction that neither instruction occurred more than twice in succession. Although the categories were presented to every subject in the same order, the order of items within categories was randomly determined for each subject. The counterbalancing variable, study list, ensured that an equal number of subjects at both retention intervals saw every category with each instruction.

The directions read to the subject before list presentation described the structure of the list and emphasized the importance of using the F instruction, since the list was long and would be presented only once. The subject was assured that F items would not be tested; he was also cautioned that the order of F and R categories was random, so that it would not be worthwhile to try to anticipate the upcoming instruction.

All subjects were given an immediate self-paced recall test on only the four buffer categories. The buffer category names were provided and the subject was asked to recall the three exemplars from each, guessing if necessary. This test was administered to help ensure that subjects in the 1-wk group would not anticipate the retention test.

This buffer test was followed either immediately or 1 wk later by three further tests, the first one being a free-recall test. Subjects were instructed to recall as many items as they could by writing down

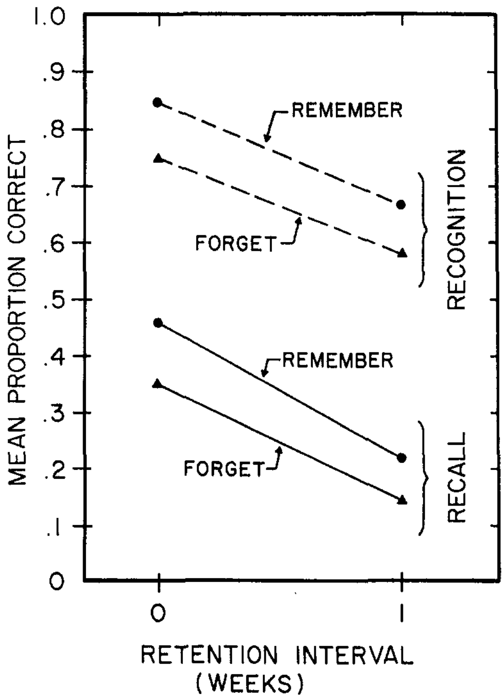

FIGURE 1. Mean proportion of responses correct in cued recall and in recognition as a function of retention interval (0 or 1 wk) and instruction (*remember* or *forget*).

the category names and as many of the three exemplars as possible for each. In addition, the names of the already-tested buffer categories were provided to prevent subjects from including them in free recall. The free-recall test was followed by the two remaining tests, category-cued recall and three-alternative forced-choice recognition of the items in each of the 16 critical categories.

All subjects received each of the tests in the order just mentioned, so that the two types of recall tests and the recognition test cannot be considered independent measures. The sequence used was chosen so that the tests with a larger retrieval component preceded those with a smaller retrieval component. The major problem in multiple testing is usually that an item that could not have been output on a given test becomes available because of a preceding test. The present order of tests was selected with the intent of minimizing possible effects of prior tests on succeeding tests.

All tests were self-paced with none requiring more than 10 min. The response sheets for both free and cued recall had 16 rows and 4 columns; however, in cued recall the first column contained the category names (in a random order different from that used in list presentation), while in free recall this column was blank to allow the subject to enter the category names.

The recognition test was conducted using a deck of 3 × 5 in. (8 × 13 cm) index cards prepared as follows. Each card consisted of 3 category exemplars typed horizontally. One exemplar was an original study item, assigned at random to a position on the card; the other two exemplars were that item's distractors. The recognition test deck was blocked by category, with the order of categories and cards within categories randomly determined for each subject. In cued recall and in recognition, the subject was required to provide three responses per category; in free recall, omissions were allowed.

Results and Discussion

In both Experiments 1 and 2, all significant statistics are reliable beyond the .001 level unless otherwise indicated; nonsignificant statistics are based on $p > .10$. Mean square errors (MS_e) are included in all cases where they could not be derived by the interested reader, i.e., those cases where exact means are not presented.

The comparability of the two retention-interval groups in proportion correct, $P(C)$, on buffer items, $F(1, 46) = 1.01$, suggests that the 0-wk group ($M = .674$) and the 1-wk group ($M = .733$) were approximately equal in acquisition.

In the retention-test phase, a free-recall test preceded the cued-recall and the recognition tests. Because of the very low absolute level of performance in this task by the 1-wk group, $P(C) < .05$, these data are not discussed further. It is worth noting, however, that R items ($M = .25$) were significantly better free-recalled than F items ($M = .15$) in the 0-wk group, $F(1, 46) = 6.73, p < .05$.

Separate 2 × 2 × 2 (retention interval × instruction × study list) analyses of variance were conducted on the cued-recall and recognition data. Figure 1 presents the mean P(C) for R items and F items as a function of retention interval for both recognition and cued recall. Turning first to the cued-recall analysis, the effect of retention interval was highly significant, $F(1, 44) = 26.27$, $MS_e = .04$, accounting for more than 35% of the between-subjects variance ($\omega^2 = .352$). The effect of instruction was also highly significant, $F(1, 44) = 29.17$, $MS_e = .01$, accounting for slightly less than 29% of the within-subjects variance ($\omega^2 = .289$). Qualifying this instruction result somewhat was the significant Study List × Instruction interaction, $F(1, 44) = 22.51$, $\omega^2 = .220$. The directed forgetting effect was very strong in one list, $t(94) = 10.22$, $MS_e = .02$, but virtually absent in the other, $t(94) = .65$. This category-specific effect of instruction has no obvious explanation, but suggests that more, different items should have been used.

Central to the present study is the Instruction × Retention Interval interaction. If the directed forgetting effect were to diminish over retention interval, this would suggest a breakdown in differentiation of R and F items in storage, implicating a selective search account. However, this critical interaction was nonsignificant ($F < 1$).

The conclusions from the recognition analysis are identical to those from the cued-recall analysis whether $P(C)$ or d' is used as the dependent variable; for brevity, only the $P(C)$ analyses are reported throughout this article. The effects of retention interval, $F(1, 44) = 35.38$, $MS_e = .02$, $\omega^2 = .414$, and of instruction, $F(1, 44) = 27.14$, $MS_e = .01$, $\omega^2 = .347$, were both highly significant. Also, as in cued recall, the Instruction × Retention Interval interaction was nonsignificant ($F < 1$).

Taken together, the recognition and cued-recall results clearly demonstrate the consistent superiority of R items over F items. The directed forgetting effect [$P(C)$ for R items − $P(C)$ for F items] was approximately 10% at both retention intervals, with the test delay simply lowering the absolute level of performance.

The remarkably consistent stability of the effect across retention interval and type of retention test is difficult to align with the selective search account for two reasons. First, although not required by selective search, the prediction of an interaction between instruction and retention interval is tempting, if it is assumed that the set-differentiating information required for selective search is lost before the item-specific information. This interaction does not appear in recall or in recognition. Second, and more crucial, is the prediction from selective search of a reduced directed forgetting effect in recognition as compared to recall. This prediction is also contradicted, leaving selective search considerably weakened as the primary mechanism underlying directed forgetting. On the other hand, the results are fully in accord with the selective rehearsal prediction of a constant directed forgetting effect due to greater initial processing of R items than of F items.

It appears, then, that R items are more available than are F items. To what initial processing difference might this be attributed? It would seem reasonable, as Woodward et al. (1973) have proposed, that items are simply maintained, or recirculated, until their instruction is presented. Subsequent to an R instruction, the R items are then elaboratively rehearsed, whereas subsequent to an F instruction, previous R items may be elaboratively rehearsed, but F-item rehearsal will be discontinued altogether. Application of this account to the present experiment suggests that F and R items would be maintained for the same length of time, and the R-item advantage would be due to the restriction of elaborative rehearsal only to those items to be remembered.

Experiment 2

The recognition findings are the key results of Experiment 1. One of the aims of Experiment 2 was to extend these results using a set of unrelated items with item-by-item instructions. Additionally, tests at 1- and 2-wk retention intervals (between subjects) were used to help extend the conclusions regarding the stability and duration of the directed forgetting effect.

Experiment 1 demonstrated that the role of selective rehearsal is central, but might not selective search also play some role in directed forgetting? Using paired associates, Epstein and Wilder (1972) have shown that when a subject is informed at test that a given item is an F item, the directed forgetting effect is eliminated. This result initially seems to favor a selective search interpretation wherein the subject has a higher probability of searching the R set than the F set. However, in their study, F instructions were given only after the entire set of items was presented (i.e., not intraserially), essentially precluding selective rehearsal. Also, the subject was informed beforehand that some F items would be tested. Together, these manipulations greatly increase the problem of demand characteristics (Orne, 1962). Although Reitman, Malin, Bjork, and Higman (1973) compensated in part for this problem, their method still required occasional tests of F items. Thus, Experiment 2 introduces a modified form of instructional cue at the time of a recognition test in an intraserial version of the directed forgetting paradigm. As in Experiment 1, there is only one extended test trial, either 1 or 2 wk after acquisition.

Method

Design. The design was a 2 × 2 × 2 × 2 × 2 factorial, with 3 between-subjects variables and 2 within-subjects variables. Two of the between-subjects variables were for purposes of counterbalancing. These were (a) the instruction with an item at study (study list) and (b) the presence or absence of an instruction cue with a particular item at test (cue set). The remaining between-subjects variable was retention interval (1 or 2 wk). The two within-subjects variables were (a) instruction during list presentation (*remember* or *forget*), and (b) presence or absence of an instructional cue at test (cued or uncued). An instructional cue presented with an item at test was always the same as the instruction presented with that item at study.

Subjects. There were 80 subjects, all University of Washington undergraduates whose participation partially fulfilled a course requirement. Ten subjects were assigned at random to each of the 8 groups defined by factorial combination of the between-subjects variables. The subjects were divided into subgroups, varying in size from 3 to 7, during both study and test.

Stimuli and apparatus. The study items and recognition test distractors were all A and AA nouns chosen from the Paivio, Yuille, and Madigan (1968) norms, without respect to imagery or meaningfulness values. All words were one- or two-syllable nouns, four to eight letters in length. For the 38 study items, first-letter frequency was restricted to a maximum of four times per letter. All study items were typed in uppercase letters and made into slides, as were the instructions. Slides were presented via a Kodak Carousel projector with an external timing device.

Two of the 76 test distractors were assigned to each study item, with the restrictions that (a) the study item and its two distractors began with different letters and (b) the two distractors were not related semantically to their study item.

There were two three-alternative forced-choice recognition tests, constructed in the following manner. The first test was given immediately after study as a decoy test in an effort to prevent the subject from anticipating the retention test. The 6 study items on this test consisted of the first 3 items and the last 3 items from the study list, which were always R items. These 6 buffer items were included at study with the dual purpose of permitting the decoy test and of reducing serial position effects on the 32 critical study items. The decoy test was constructed by assigning 1 of the 6 buffer items to each of the six rows on a 6 × 3 response sheet, with the column of each buffer item randomized. The remaining two positions in each row contained the distractors assigned to that buffer item.

The second recognition test was given either 1 or 2 wk after study. On this test, the 32 critical items were arranged on a response sheet similiar to the buffer-test response sheet. However, there was also a fourth column that was either blank for a given row (uncued), or contained an R or an F informing the subject of the original study instruction paired with the study item in that row (cued). For each subject, half of the R items and half of the F items were cued; the other half were not cued. The counterbalancing variable cue set ensured that each item was equally often R-cued and F-cued and that each item was cued as often as it was uncued.

Procedure. The 38 items were shown to every subject in the same order, while instructions were counterbalanced by the study-list variable. Presentation rate was 3 sec for each item and 3 sec for its instruction, which immediately followed the item. The first three and last three buffer items were always R items. Of the 32 critical items, half received R instructions and half received F instructions, with the restriction that neither instruction could occur more than three times successively.

Directions read to the subjects before list presentation emphasized the usefulness of the F instruction, since the list was long and would be presented only once. The subjects were also told that F items would not be tested and were cautioned concerning the random order of the F and R instructions. Following list presentation, subjects were given the decoy test, which was self-paced, although all subjects completed it in less than 2 min. The subjects were not warned about the retention test.

The subjects returned either 1 or 2 wk later for the test on the 32 critical items. This test was also self-paced, no subject requiring longer than 12 min to complete it. The subjects were instructed to circle the correct word in each row, omitting no rows. They were told that when an R or an F appeared to the right of a row, this would indicate the instruction given with that item at study and might help them in choosing the correct item. They were also told that when there was a blank to the right of the row, they were to fill in an R or an F to indicate which instruction they thought had been paired during study with the item chosen. Response sheets were checked for omissions before the subjects were dismissed.

Results and Discussion

An analysis of variance was conducted on the $P(C)$ data based on the 32 critical items. (The analysis performed using d' produced essentially the same conclusions, with the two exceptions noted below.) The decline in $P(C)$ from 1 wk ($M = .61$) to 2 wk ($M = .56$) was marginally significant, $F(1, 72) = 2.87$, $.10 > p > .05$; this decline was significant using d' ($p < .05$). The effect of instruction was highly significant, $F(1, 72) = 72.17$, $MS_e = .02$, while the Instruction × Retention Interval interaction did not approach significance ($F<1$). These results indicate that the directed forgetting effect obtained in Experiment 1 is reliable in an altogether different procedure and remains stable across retention intervals as long as 2 wk. The heavy lines in Figure 2 show mean $P(C)$ in recognition of R items and F items as a function of retention interval. Once again, the results support the selective rehearsal account over the selective search account, since the directed forgetting effect is approximately 15% at both retention intervals in Experiment 2.

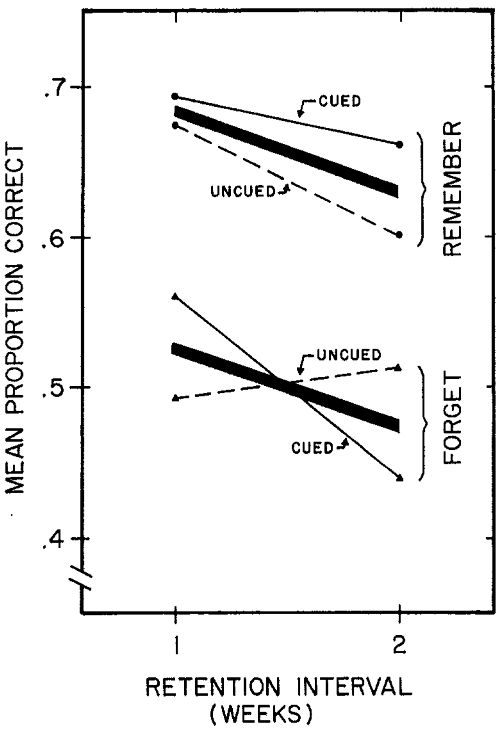

FIGURE 2. Mean proportion of responses correct in recognition as a function of retention interval (1 or 2 wk). (Heavy lines represent instruction [*remember* or *forget*] collapsed over the cue variable [cued or uncued]).

TABLE 1

PROPORTION OF UNCUED ITEMS ASSIGNED AN ACCURATE INSTRUCTION LABEL AT TEST AS A FUNCTION OF RETENTION INTERVAL

Retention interval	Proportion labeled accurately		
	"F"/F	"R"/R	Overall
1 wk	.56	.65	.61
2 wk	.52	.67	.60

Note. The proportion of all items labeled "R" (response bias) was .55 at 1 wk and .58 at 2 wk. Abbreviations: F, forget; R, remember.

The effect of instructional cue presented during recognition testing was not significant ($F < 1$), nor was the Cue × Instruction interaction, $F(1, 72) = 2.52$, $MS_e = .03$, indicating that cuing F items does not improve their recognition relative to R items. This contradicts the prediction from selective search that the directed forgetting effect is reduced or eliminated for those items where, during test, the subject is informed of the related study instruction (cf. Epstein & Wilder, 1972).

Although the present results do not support the notion that subjects use a selective search strategy, it seems that subjects do store instruction information with items at the time of study. Table 1 presents the proportions of uncued items that were assigned the correct instruction label (at the time of testing) as a function of retention interval. The proportions shown ill the last column are estimates of overan accuracy on both F and R labels. (These accuracy estimates should be relatively free of response bias, which acts in opposite directions on the F and R items.) Overall labeling accuracy was significantly greater than chance (.50) at both the 1-wk, $t(39) = 5.79$, and the 2-wk, $t(39) = 5.00$, retention

TABLE 2

Conditional Probabilities Relating Item Recognition to Recall of Study Instructions

Conditional probability (P)	Chance	Observed	z	p
$P(C_i \mid C_r)$	.50	.65	8.20	<.001
$P(C_i \mid E_r)$	.50	.53	1.38	>.05
$P(C_r \mid C_i)$	.33	.63	17.65	<.001
$P(C_r \mid E_i)$	.33	.51	8.61	<.001

Note. The C represents a correct response; E represents an incorrect response. The subscript r represents item recognition performance; the subscript i represents instruction recall performance. The test statistic comes from the normal approximation to the binomial.

intervals. Furthermore, the decline in accuracy with retention interval was nonsignificant, $t(78) = .37$. Of greater interest are the conditional probabilities presented in Table 2, relating recognition of an item to recall of its instruction [e.g., $P(C_r \mid E_i)$ represents the proportion of items correctly recognized (r) given that their associated study instruction (i) was incorrectly recalled]. These results suggest that the subject does have some information in storage concerning instructions, but that this information is retrievable only for items that are recognized. This is demonstrated by the significant $P(C_i \mid C_r)$ versus the nonsignificant $P(C_i \mid E_r)$. On the other hand, the subject does not require information concerning the instruction associated with an item at study in order to recognize that item. This is shown by the $P(C_r \mid E_i)$ being significantly above chance.

Finally, there were several significant interactions in the analysis of variance that are not pertinent to the issues being discussed and will therefore simply be briefly stated. These include the Study List × Cue and Cue Set × Cue interactions, both merely indicating order effects due to counterbalancing. The Study List × Instruction interaction was also significant, $F(1, 72) = 16.57$, $MS_e = .02$. However, simple effects t tests revealed that the directed forgetting effect was highly reliable for both lists. Two second-order interactions were significant: Retention Interval × Cue Set × Cue [$p < .05$ in $P(C)$, but $p > .05$ in d'], and, as is clear from the lighter lines in Figure 2, Retention Interval × Cue × Instruction ($p < .05$). No obvious interpretation of either is apparent, and none of the first-order interactions of the retention interval, cue, and instruction variables was significant.

General Conclusions

Taken together, the results of these two experiments suggest an account for the directed forgetting effect. The proposed account subsumes the four key findings of the present study: (a) the comparability of the directed forgetting effect in recall and in recognition, (b) the constancy of the effect across relatively long retention intervals, (c) the failure of study instructions presented at test to reduce the effect, and (d) the retrievability of instruction information for those items that are recognized but not for those items that are not recognized.

The directed forgetting effect is assumed to be due primarily to rehearsal being limited to R items (i.e., the selective rehearsal account). In line with Woodward et al. (1973), it is suggested that an item is simply maintained by rote rehearsal until the related instruction is presented, at which time elaborative rehearsal occurs if the item is an R item. Since the directed forgetting effect stems from these initial processing differences, it should persist unchanged across longer retention intervals. This stability and duration of the effect is clearly obtained in the present two experiments, whether the measure of retention is recall or recognition. As pointed out, the presence of the directed forgetting effect in recognition contradicts a purely selective search account of directed forgetting, wherein the effect is assumed to be due to a higher post-input probability of searching the R set than the F set; furthermore, Experiment 2 demonstrates that the directed forgetting effect persists in recognition even given provision of the appropriate study instructions as cues at test, a finding also contradictory to selective search.

Since the advantage of R items over F items is apparently established during initial processing and persists thereafter, directed forgetting may be one way in which subjects selectively process inputs to avoid overloading their processing capabilities. However, the associated study instruction also seems to be encoded, since subjects are able to retrieve this type of information when the item is recognized.

Thus, two major implications of the present study are: (a) Selective rehearsal during initial processing produces a long-term advantage of R items over F items regardless of whether the performance measure is cued recall or recognition, and (b) instruction information, although not used directly for selective search, is stored as part of the representation for items encoded in memory.

REFERENCES

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 1969, *80*(3, Pt. 2).

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.: V. H. Winston, 1972.

Bjork, R. A., Laberge, D., & Legrand, R. The modification of short-term memory through instructions to forget. *Psychonomic Science*, 1968, *10*, 55–56.

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

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.

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

Elmes, D. G., Adams, C., & Roediger, H. L. Cued forgetting in short-term memory: Response selection. *Journal of Experimental Psychology*, 1970, *86*, 103–107.

Elmes, D. G., & Wilkinson, W. C. Cued forgetting in free recall: Grouping on the basis of relevance and category membership. *Journal of Experimental Psychology*, 1971, *87*, 438–440.

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

Epstein, W., Massaro, D. W., & Wilder, L. Selective search in directed forgetting. *Journal of Experimental Psychology*, 1972, *94*, 18–24.

Epstein, W., & Wilder, L. Searching for to-be-forgotten material in a directed forgetting task. *Journal of Experimental Psychology*, 1972, *95*, 349–357.

Kintsch, W. *Learning, memory, and conceptual processes*. New York: Wiley, 1970.

Murdock, B. B., Jr. Modality effects in short-term memory: Storage or retrieval? *Journal of Experimental Psychology*, 1968, *77*, 79–86.

Muther, W. S. Erasure or partitioning in short-term memory. *Psychonomic Science*, 1965, *3*, 429–430.

Orne, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, *17*, 776–783.

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

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.

Shebilske, W., Wilder, L., & Epstein, W. Forget instructions: The effect of selective rehearsal and categorical distinctiveness. *Journal of Experimental Psychology*, 1971, *89*, 372–378.

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. Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 608–617.

(Received August 9, 1974;
revision received November 11, 1974)