Journal of
Experimental Psychology

Vol. 101, No. 1 November 1973

LEARNING AND INTERFERENCE EFFECTS IN SHORT-TERM MEMORY [1]

LYNN HASHER [2]
Temple University

JUDITH GOGGIN
University of Texas at El Paso

DONALD A. RILEY
University of California, Berkeley

The present experiment used essentially a transfer-of-training paradigm to study the fate of individual items in the Brown–Peterson paradigm. The *S*s first received 3 tests on triads that came from the same taxonomic category. This was followed by 3 tests on triads from a new category. In the second half of the experiment *S*s received either a third category (Release), the first-presented category with the same items (Same), the first-presented category with different items (Different), or they continued with items from the second category (Control). "Learning" effects were found in Condition Same, "interference" effects in Condition Different, and release from proactive inhibition effects in Condition Release. Together with evidence from a final recall test, these data present strong evidence for specific items being stored and retained across successive Brown–Peterson tests, despite the typically observed massive interference.

The question of whether there is a short-term memory independent of a long-term memory can most convincingly be answered by a demonstration that the 2 memories obey different laws of acquisition and retention. Although it has been proposed that information in the short-term store is encoded phonemically while information in the long-store is encoded semantically, recent research has shown that both phonemic and semantic encoding occur in both the short- and long-term memories (Glanzer, Koppenaal, & Nelson, 1972; Nelson & Rothbart, 1972; Shulman, 1972). While a rationalization of these overlapping encoding effects that is consistent with the assumption of 2 memory processes may be found (e.g., Glanzer et al.), such results have at least eliminated the simplicity of the original assumption that the 2 stores differ in terms of the kinds of information that they process.

With regard to retention, it has been argued that memory traces in the short-term store decay while those in the long-term store are interfered with or suffer a loss of accessibility (Tulving & Pearlstone, 1966). Attempts to demonstrate decay processes (e.g., Peterson & Peterson, 1959) have been hampered by the fact that classical interference mechanisms have been shown to operate in tasks such as the Brown–Peterson paradigm (Brown, 1958; Peterson & Peterson), which presumably tap short-term memory (Keppel & Underwood, 1962). Furthermore, indirect attempts which might have shown decay in

[1] This research was supported by the Committee of Research and the Institute of Human Learning, University of California, Berkeley. The authors acknowledge with thanks the advice of Geoffrey Keppel and Leo Postman.

[2] Requests for reprints should be sent to Lynn Hasher, Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122.

the short-term store have actually found evidence for long-term storage (Hebb, 1961) in a "short-term" task. Thus, strings of items that are repeated after lags filled with the immediate recall of other strings of items have a higher probability of being recalled than do nonrepeated strings.

While the pure decay of items has no support in the literature, there is some recent evidence (Cermak, 1970; Kincaid & Wickens, 1970) that suggests that the proactive interference (PI) that occurs across successive Brown–Peterson trials may itself decay in time. The level of release from PI, or improvement in performance, was found to be a function of the length of the delay introduced between 2 successive test trials. An alternative interpretation of this effect, of course, is that differentiation, a process that in long-term memory tasks is known to reduce interference (Dallett & Wilcox, 1968; Underwood & Ekstrand, 1967; Underwood & Freund, 1968), was operating here in the guise of a distinctive time tag (Yntema & Trask, 1963) that enabled the superior retrieval of the last-presented items.

The present experiment uses a variation of the repetition design of Hebb (1961) to ask the question of the fate of individual items in the Brown–Peterson task at a time when, if interference does indeed dissipate, the old items that produce interference should no longer be available to compete with new items. Such a goal was achieved with a rather complex experimental design that had 4 main types of conditions: (*a*) Release, in which the category of the to-be-remembered material was changed; (*b*) Control, in which under comparable circumstances a category change did not occur; (*c*) Same, in which following an initial series and a release series, *S*s were returned to the initial category and were tested again with the identical set; and (*d*) Different, in which *S*s were treated identically to those in the Same condition except that the return to the initial category involved different, rather than the same, items.

The conditions of most significance for the question of the fate of individual items were Conditions Same and Different. In each, were there no record of the first-presented category or items, a return to old categories should be no different from a change to an entirely new category, and one would predict performance comparable to that obtained under the Release condition. If, however, specific items from the first category are still available after the interpolation of a second category, a repetition or learning effect would obtain under Condition Same and, likewise, an interference effect under Condition Different. A further measurement of specific-item retention was obtained in a final recall test given to *S*s after 12 Brown–Peterson trials.

Method

Procedure. The main part of the experiment comprised 12 Brown–Peterson test trials divided into 4 successive blocks of 3 trials each. On each test 3 words were presented for *S* to recall. Prior to the first experimental trial, *S* received 2 practice trials to familiarize him with the task. The practice trials were procedurally the same as experimental trials but used symbols instead of words as the to-be-remembered materials. Each test trial was identical. First, the word READY appeared for 2 sec. This was followed by a 1.5-sec. presentation of 3 words printed diagonally across the slide; *S* read the words aloud. A randomly selected 3-digit number then appeared, from which *S* was required to rapidly count backward by 3s. After 15 sec. of number counting, a question mark appeared for 10 sec. This was the signal for *S* to recall the 3 words he had just seen. The next test trial followed immediately. All materials were presented visually by a Carousel projector with exposure duration regulated by a Wollensak Slide-Sync tape recorder. At the beginning of the experiment, *S* was instructed to recall the words in order, if possible.

Following the 12 tests, an unpaced written recall test was administered, with *S* instructed to recall as many words as he could. Half of the *S*s in each condition were given no other information at recall; the remaining *S*s were cued with the appropriate category name(s) for the items they had seen during the 12 test trials. These names were written at the top of each *S*'s recall sheet.

Design. The 8 conditions in the experiment are outlined in Table 1. The letter C in the body of the table stands for the word *category*. The numerical subscript indicates the number of different categories presented to that point in the experiment. For the purpose of congruence with the description of the results, the design will be explained in the following order: Control and Release conditions (Conditions 1–4), then Same and Different conditions (Conditions 5–8).

All conditions were treated alike in the first block of trials, with the items for any one *S* all coming from the same category (C_1). Condition 1 continued with instances from C_1 in the second block; all other *S*s were given instances from a new category (C_2). Condition 1 then provides a control for whatever release from PI is observed with a change to a new category. In Block 3, Conditions 3 and 4 change to a third new category (C_3), with Condition 2 providing the no-change control. In Block 4, a change to a fourth new category (C_4) by Condition 4 can be compared to Condition 3, now a no-change control. Thus, the pattern of Conditions 2-4 is such that for every block, one or more conditions are changed to a new encoding category. For each of these changes, there is one condition which has been treated identically up to that point that serves as a no-change control.

Different types of experimental manipulations were introduced to the remaining conditions in the final 2 blocks. In Table 1, the letters s and d indicate, for those conditions returning to an old category, whether the words are the same or different, respectively, from those originally presented.

First, consider Conditions 5 and 6. In Block 3, C_1 was presented again to both groups, with the old instances occurring once again in the same order on each test. For example, if the triad LETTUCE, ONION, MILK had appeared on Test 1, it would again appear on Test 7. In Block 4, treatment of these 2 conditions diverged. The item-repetition pattern was continued for Condition 6, with C_2 being presented again. Thus, for Condition 6 the second half of the experiment was an exact replicate of the first. A new category (C_3) was presented to Condition 5 in the last block of trials.

The final 2 groups (Conditions 7 and 8) also received a repetition of C_1 in the third block, but the specific items presented were new. For example, if the triad LETTUCE, ONION, MILK were presented on Test 1, these conditions might receive HAM, PRUNE, COFFEE on Test 7. Condition 7 continued with C_1 in the fourth block, thus receiving no change in category. Condition 8, on the other hand, received new instances of C_2 in Block 4.

Materials. Four categories of words were used: foods, animals, clothing, and body parts. Each category contained 36 instances divided into 12 subsets of 3 words each. Within each condition, each of the 48 triads appeared once in each of the 12 test positions. The order of presentation of the categories was also counterbalanced across each condition.

Subjects. There were 48 *S*s in each condition (total $N = 384$). The *S*s were students and staff at the University of California, Berkeley. Ten *S*s were discarded because of equipment failure or *E* error. An additional 19 *S*s were discarded because they failed to read a triad out loud as it was presented.

TABLE 1

EXPERIMENTAL DESIGN

Condition	Sequence of categories			
	Block 1	Block 2	Block 3	Block 4
1	C_1	C_1	C_1	C_1
2	C_1	C_2	C_2	C_2
3	C_1	C_2	C_3	C_3
4	C_1	C_2	C_3	C_4
5	C_1	C_2	C_{1s}	C_3
6	C_1	C_2	C_{1s}	C_{2s}
7	C_1	C_2	C_{1d}	C_{1d}
8	C_1	C_2	C_{1d}	C_{2d}

Note. C = category; numerical subscripts indicate the number of different categories presented to that point; letter subscripts indicate, for those condition returning to the old category, whether the words are the same (s) as or different (d) from those originally presented. Conditions 1-4 are Control and Release conditions, 5-8 are Same and Different conditions.

RESULTS

The results will be discussed in 6 sections. The first section deals with comparability among the different experimental conditions and with the replication of basic phenomena. The succeeding sections are concerned with repeated release, the effects of specific-item repetition, the persistence of interference, intrusion analyses, and terminal free recall, in that order. Protocols were scored by assigning 1 point for each item correctly recalled regardless of position. A stringent score that awarded a point only if an item were given in its correct position was also tallied. Although analyses were done on both scores, only those on the lenient measure will be reported. In general, the stringent scoring revealed the same effects as did the lenient scoring.

Replication of basic phenomena. The first 6 test trials essentially replicated an earlier study (Loess, 1967) that used taxonomic categories as materials in the Brown-Peterson paradigm. These data are shown in Panel 1 of Figure 1. To facilitate comparisons of performance in Block 2, the conditions that become Release conditions in the second block, that is, Conditions 2-8, are grouped together; Condition 1, the control for release effects, is plotted separately.

On the first 3 trials, the 8 conditions were treated alike. A repeated measures analysis of variance showed neither a main effect for conditions nor an interaction between conditions and test trials. This lack of difference among the groups in Block

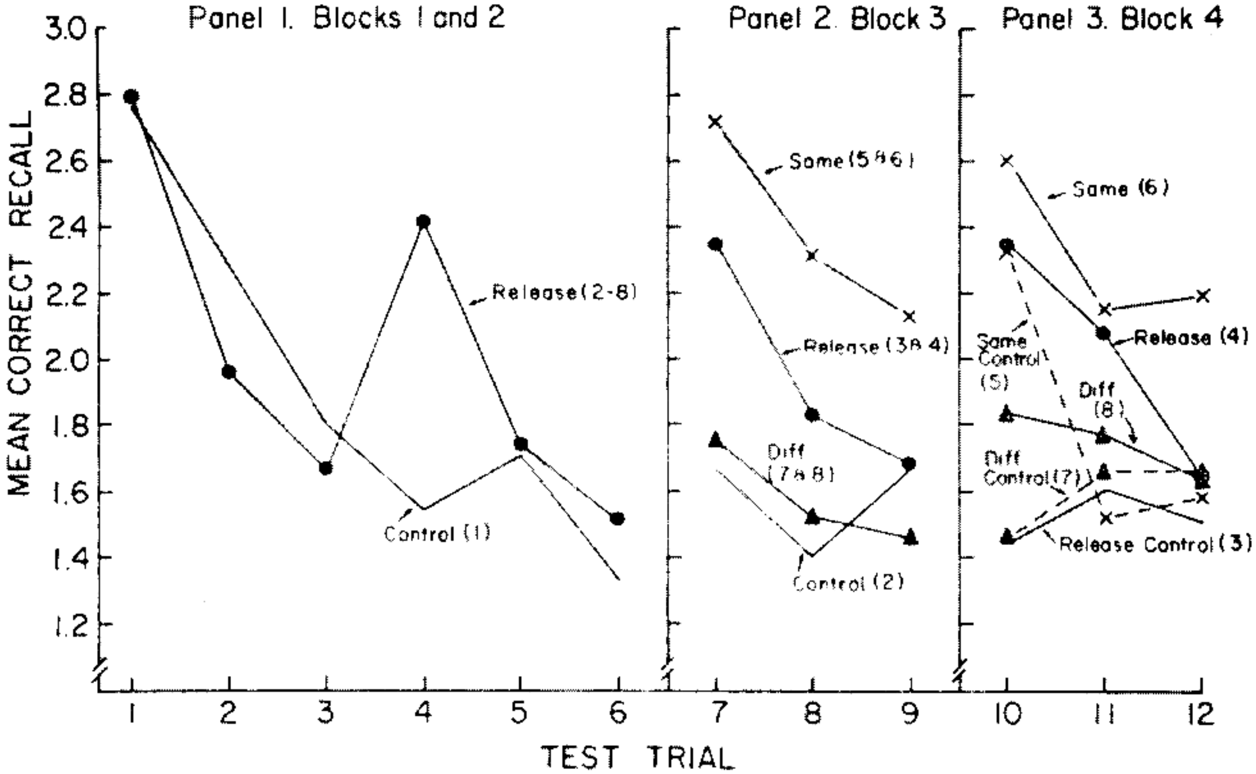

FIG. 1. Mean correct recall on Blocks 1–4 of the Brown–Peterson task.

1 indicates that they were equated in ability. Across Tests 1–3, there was a pronounced decline in the level of recall, producing a significant main effect of tests, F (2, 752) = 206.51, $p < .001$. Thus, as in previous research (cf. Wickens, 1970), PI built up rapidly with only a few prior items.

On Test 4, a new category was introduced to all *S*s except those in Condition 1, who continued with instances of the original category. A category change produced a marked increase in recall on Test 4. The difference between the performance of Control *S*s and the average performance of Release *S*s on this test was reliable, F (1, 94) = 20.84, $p < .001$.[3] An overall analysis on Tests 4–6 comparing Condition 1 with the combined 7 Release conditions showed that performance of the latter *S*s was superior, F (1, 94) = 8.30, $p < .01$. Although recall generally declined over the second block of trials, F (2, 188) = 7.85, $p < .001$, experimental and control *S*s behaved quite differently. Recall of the Release conditions decreased across tests, which is evidence for the buildup of PI in the new category; recall of the Control condition remained poor across tests. The interaction between the control and the experimental conditions across test trials was significant, F (2, 188) = 5.07, $p < .01$.

Repeated release. Each time a new category was introduced, release from PI occurred again, and it occurred each time to approximately the same extent. This can be seen in Block 3 (Panel 2, Figure 1) by comparing the curves labeled Release and Control. Release data derive from *S*s in Conditions 3 and 4, who were introduced to a new category on Test 7; Control data derive from *S*s in Condition 2, who had no category change at this point. In Block 4 (see Panel 3), *S*s in Condition 4, who were presented a fourth new category

[3] Because of the widely discrepant *n*s in several of the analyses performed in this experiment, we judged that use of 2 weighted-means analyses of variance was inappropriate. Instead, in each of the analyses involved we computed the sums of squares using means that represented the disparate groups as having equal *n*s. In this particular instance the Release conditions were treated as a single group with 48 *S*s. Analyses were then done on the mean values with the degrees of freedom concomitantly reduced. The tests of significance were, then, extremely conservative. This same plan was used throughout the experiment, as appropriate.

on Test 10, provide the Release scores, while *S*s in Condition 3 provide the Control scores.

Planned comparisons on Tests 7 and 10 between these Release and Control conditions again showed that there was significant release from interference following a category change, *F*s (1, 376) = 14.80 and 21.01, respectively, $ps < .001$. Across tests within each block, overall recall of Release *S*s was higher than that of Control *S*s, yielding F (1, 94) = 9.06, $p < .01$, for Block 3, and F (1, 94) = 16.68, $p < .001$, for Block 4. Finally, conditions and tests reliably interacted, *F*s (2, 188) = 3.29 and 3.93 for Blocks 3 and 4, respectively, $ps < .05$. This indicates that while recall declined over tests in the Release conditions, no further decrement in performance was observed under Control conditions.

In summary, changes in the encoding category of the materials repeatedly and reliably improved recall performance on the trial on which the change was introduced. Thereafter, with subsequent tests on the new category, a drop in recall was obtained.

It should be pointed out here that data from Condition 1 and from Conditions 1 and 2 are omitted from Panels 2 and 3, respectively, because in each case their principal purpose was to serve as a baseline against which to compare performance in earlier blocks. However, after a second or third test following introduction to a new category, recall in these conditions did not change appreciably. It appears that PI builds up to a maximum after only a few tests, and, if there is no alteration of the stimuli, remains stable thereafter.

Repetition effects. Repetition effects may be observed in those conditions (5 and 6) that return to the same instances of an earlier category. In Block 3, the 2 Same conditions (5 and 6) are compared with the 2 Release conditions (3 and 4). This pattern of conditions is replicated in the fourth block: Same *S*s (Condition 6) are those who repeat the items from Block 2; Control *S*s (Condition 5), labeled Same Control in Panel 3, were changed to a new category after 9 test trials that were identical to those of *S*s in Condition 6. The comparisons made in Blocks 3 and 4 are between conditions that have studied a category either once or twice.

First, consider the Same conditions in Block 3. Repetition of items had a facilitative effect over and above that found with a change to a new category. The superiority of Same over Release conditions was confirmed both by a planned comparison on Test 7, F (1, 376) = 7.62, $p < .01$, and by a comparison across Tests 7-9, F (1, 190) = 26.70, $p < .001$. The latter analysis also showed that recall dropped over successive tests in Block 3, F (2, 380) = 26.95, $p < .001$, and that conditions did not interact with tests. The Same and Release curves in Panel 2 demonstrate these effects clearly. Note that interference increased across tests, even under Same conditions where old items were relearned. The advantage of repeating old instances over learning new ones appears to be constant over the 3 trials in Block 3.

The item repetition effects in the last block are somewhat less clear than are those in Block 3, although they remain in the same direction. Panel 3 shows the Same condition to have an advantage over the Control condition. While the effect was not significant on Trial 10, it was when all 3 trials of the fourth block were considered, F (1, 94) = 13.74, $p < .001$. Test effects in Block 4 mirrored those in Block 3; there was a reliable decline in recall across test trials and no interaction between conditions and tests. Thus, in both blocks, item repetition was effective in elevating performance above Release conditions, clear evidence that under these conditions *S*s remembered the specific items they had been exposed to and were not merely showing the release associated with a new category.

Persistence of interference. The final conditions to be considered are the Different conditions. In Block 3, Conditions 7 and 8 received different instances of the category used in Block 1. These 2 conditions are combined in the curve labeled Different in Panel 2. Because of large differences between Different and Release conditions, the statistical comparisons reported are be-

TABLE 2

FREQUENCY AND SOURCES OF INTRUSIONS

Intrusions	Conditions[a]							
	1	2	3	4	5	6	7	8
Total number	176	191	146	151	131	148	159	186
Number per *S*	3.67	3.98	3.05	3.15	2.73	3.09	3.32	3.88
Proportion from each source								
Intralist	.88	.87	.80	.83	.85	.86	.91	.84
Same category	1.00	.99	.95	.94	.96	.96	.98	.94
Test $n-1$	.46	.71	.75	.86	.89	.71	.71	.62
Test $n-2$	.19	.12	.14	.14	.10	.20	.10	.15
Tests $n-3$ or more	.35	.17	.12	.00	.01	.10	.18	.22

[a] Conditions 1–4 are Control and Release conditions; 5–8 are Same and Different conditions.

tween the Different conditions and those Control conditions that continued in the category from the prior block of 3 trials. In Block 3, the appropriate comparison is made between the Different conditions and Condition 2. In Block 4, the appropriate comparison is made between Conditions 7 and 8, with the former continuing to receive exemplars from the category used in Block 3, and the latter receiving exemplars from the category previously used in Block 2.

In Block 3 and more so in Block 4, there is a suggestion of an advantage shown by the Different condition over the Control condition. In Block 3, performance of the Different conditions was not significantly better than that of the Control condition on Test 7. Across Tests 7–9, neither conditions nor tests was a significant factor and they did not interact with each other. While discrepancy between Different and Control conditions was somewhat larger on Test 10 than on Test 7, it did not quite reach significance, F (1, 376) = 3.52, $p >$ .05. The pattern of results over all tests in Block 4 was the same as in Block 3, with the analysis yielding no significant effects.

These results are rather ambiguous. It is possible that the superiority of the Different conditions on the first trial in both blocks may be a repeatable effect despite the lack of statistical significance. Nonetheless, the effect is small, and it is clear that in comparison with Release conditions, substantial PI occurs despite the intervening trials on a different category.

Performance shown by the Same and Different conditions leaves us with 2 conclusions: (*a*) a second presentation of old items results in improved performance, relative to a standard release condition; (*b*) a second presentation of an old category but with new instances results in poorer performance than does a change to a new category. Further, the level of performance of the Different conditions is approximately equivalent to that under conditions where no category change was introduced. Both results argue strongly that specific items are stored during Brown–Peterson trials and that these are available to *S* even after the interpolation of material from a different encoding class.

Intrusion analyses. The frequency of intrusions and a breakdown of their sources are shown in Table 2. For all conditions, the major source of intrusion errors was prior items in the list. Furthermore, of the prior items, those most likely to intrude shared the same category as the to-be-recalled items; approximately 96% of all intralist intrusions were from the same category. In addition, about 71% of these within-category intrusions came from the immediately preceding test trial.

Still earlier test-trial items appeared as intrusions in Conditions 1, 2, 3, 7, and 8. For Condition 1, all items in the list come from the same category, and consequently, all are reasonable alternatives for recall on any particular trial. Conditions 2 and 3 are controls that have more than 1 block of test trials on the same category. Conditions 7 and 8 are those that return to a

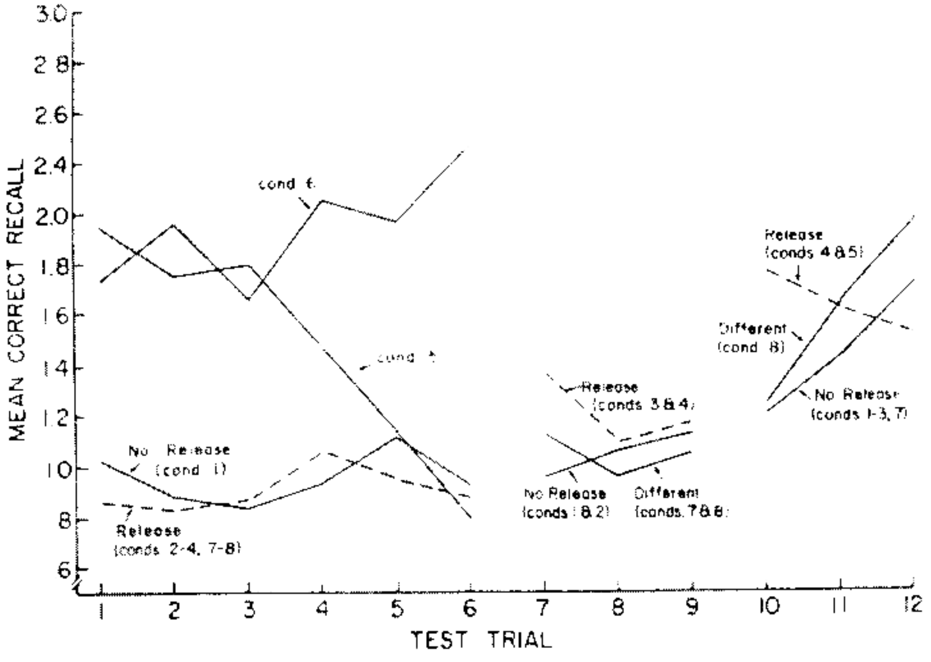

FIG. 2. Mean correct recall on final test as a function of presentation order.

previous category but with different instances. For these conditions, items from prior blocks constitute within-category alternatives for recall. For all other conditions, only the immediately preceding tests offer appropriate alternatives. That items from Blocks 1 and 2 so frequently occur as intrusions during Blocks 3 and 4 for the Different conditions suggests that earlier items are indeed free to compete at recall if appropriate to the current encoding. Thus, under Different conditions, *Ss* tend to behave as if there had never been any release category; earlier, identically coded items are apparently available to compete with later, similarly encoded items.

Terminal free recall. At the end of Block 4, all *Ss* were given an unpaced written recall test. Half of the *Ss* were asked to recall all of the words in the experiment; the other half were given appropriate category names, whose purpose was to cue recall. Because conditions differed in the total number of words presented, terminal recall was converted to a proportion of the total possible for each condition. An analysis of variance showed that only the difference among conditions was significant, $F(7, 368) = 53.00$, $p < .001$. Cuing had no effect, nor did it interact with conditions.

The data are shown in Figure 2. Four facts require mention: First, there is a clear recency effect, with items from the last block being recalled better than items from previous blocks. Second, while recall of the release triad was not uniformly better than recall of other items, it does appear to be somewhat elevated in comparison with comparable triads in the Control conditions. Thus, items which are "released from PI" tend to be better off in a terminal recall test than are other items, but the magnitude of the effect is reduced relative to performance in the Brown-Peterson task. This beneficial effect in final recall may simply reflect the fact that these items tend to be recalled in the short-term memory task, thus gaining an additional rehearsal increment relative to the nonrecalled items. Third, the conditions (5 and 6) in which specific items were repeated were superior to all other conditions. Fourth, and perhaps most important, overall recall in all other conditions was essentially the same. There is no evidence in these data for a difference in recall between *Ss* who had items in 4 categories as opposed to *Ss* who had items in only 1 category. Apparently the forgetting of items from the primacy portion of the list is not affected by the similarity of the interfering material.

A conditional error analysis of the 6 conditions that had no repeated items showed that the probability of an item being produced in the free-recall test, given that it was recalled in the Brown–Peterson task, ranged from .50 to .53. Thus, once an item is recalled, the likelihood of its being stored in long-term memory does not depend upon the experimental condition.[4] The probability of producing an item in free recall, given that it was not recalled earlier, did, however, vary with condition. This probability for Condition 4, which received 4 categories, was .11, while for Condition 1, which received only 1 category, the probability was .20. The other conditions ranged in between, in an order approximating the amount of interference presumed to be present. Assuming that this is not the product of differential successful guessing rates, it suggests that when interference is heavy, available items may not be given under the demanding task conditions of the Brown–Peterson paradigm, but can be recalled when sufficient time is allowed.

Discussion

In a design analogous to a transfer-of-training task, independent groups of *Ss* were introduced in the second half of this experiment to material that was unrelated to prior learning (Release conditions), to a repetition of prior learning (Same conditions), or to a repetition of the encoding class of prior learning, but with new instances (Different conditions). The results argue strongly that there is long-term retention of specific words in the Brown–Peterson paradigm. The Same condition showed a learning effect; items seen a second time were recalled better than new ones. The Different condition showed an interference effect; new items from an old category were more poorly recalled than new items from a new category. In fact, performance under the Different condition was close to being equivalent to that under the Control conditions that did not have a category change.

The procedure used in Blocks 1–3 in the present experiment may be conceived of as a retroactive interference (RI) paradigm if one considers just the Same and Different conditions. Under these conditions *Ss* have an original learning task (items from C_1), an interpolated learning task (items from C_2), and, in Block 3, a test for retention of original learning. While appropriate rest controls were not included in this experiment and the test for retroaction was a relearning task (reexposure to C_1), viewing the task in this light gives us an opportunity to understand the high levels of availability of C_1 items after an interpolated learning task. Tulving and Psotka (1971) have shown that the amount of RI typically observed in a list-learning situation can be substantially reduced if *S* is provided with retrieval cues stored at the time List 1 was acquired. Precisely such a situation prevailed in our experiment for the Same and Different conditions. When called upon in Block 3 to, if you will, demonstrate retention of Block 1 learning, these *Ss* were provided with a powerful retrieval cue, the encoding category used during Block 1. In one situation, Same, the high level of recall enabled *Ss* to show a "learning" effect; in the other, Different, an "interference" effect.

On the Brown–Peterson task, performance of the Same and Different conditions showed high levels of recall of individual items. In addition, performance on the terminal recall test also showed considerable memory for specific items. While recall was higher for twice-presented words, 35%–40% of all once-presented items were retained. Since cuing had no effect on amount recalled, it also appears that *Ss* had available categorical information about the items that had been presented.

The present results may be viewed as contradicting results from studies that have shown a temporally dependent release from PI. In this study, *Ss* in the Different condition returned to an old category following the interpolation of another category, and these *Ss* performed at a level equivalent to *Ss* who had never had a category change. The return to the old category occurred after an interval of time which, had it been unfilled with a second category, probably would have itself resulted in release from PI. For instance, at a 90-sec. lag, Cermak (1970) found a 90% recovery relative to the initial buildup of PI; Kincaid and Wickens (1970) found approximately 62%;

[4] No control has been exercised over the different guessing levels possible among those conditions that did not receive repeated items. Presumably, *Ss* in Condition 1 may have simply exhausted their store of, for instance, food words. The *Ss* with 2 or more categories may have proceeded through their categorical vocabularies, though they presumably needed to be more selective in their output than *Ss* in Condition 1.