No evidence of test priming between solving anagrams and completing word fragments

MICHAEL R. POLSTER and EUGENE WINOGRAD
Emory University, Atlanta, Georgia

In this paper, we report two experiments on priming or transfer effects between two implicit memory tasks—solving anagrams and completing fragments. Substantial priming effects were found only when the solution words were primed as well, either by the experimenter or by means of the subject's exposure to the item through having solved it. No evidence for test priming was found in the absence of repetition priming. Item difficulty across the tasks was highly correlated.

In presenting evidence for the existence of memory systems, Tulving (1985) emphasizes the stochastic independence found between memory tasks presumed to reflect different systems. For example, Tulving, Schacter, and Stark (1982) found independence between word-fragment completion and recognition memory. In the case of the fragment-completion task, memory is inferred from enhanced performance, or priming, following exposure to the material. This sort of task is classified as implicit, because it does not require conscious recollection of a previous episode to ensure successful performance, and it may be contrasted with tests of explicit memory such as free recall or recognition memory, which require conscious recollection of a previous episode. The kinds of memory dissociations that abound in this literature have polarized theorists into two camps: (1) those, like Tulving, who favor molar taxonomies of memory into two, three, or more systems; and (2) others, such as Jacoby (1983), Kolers, and Roediger (1984), and Roediger and Weldon (1987), who favor a more molecular level of analysis, reflected in the descriptive name of *transfer-appropriate processing*. According to this approach, performance depends on the number of operations overlapping between the tasks at encoding and retrieval.

In the present paper, we utilize a transfer design to analyze the relationship between two implicit memory tasks—word-fragment completion and anagram solution. A similar approach has been taken by Witherspoon and Moscovitch (1989), who studied priming effects in fragment completion and tachistoscopic identification. They found stochastic independence between these two tasks when tachistoscopic identification was the first task presented.

The tasks used here, anagrams and fragment completion, were chosen from the implicit memory tasks studied so far, because of their perceived similarity. Both may be generally described as involving retrieval of the correct English word on the basis of letter clues under conditions of ample time to work with the letter data. The experiments reported here were undertaken to gain some understanding of the processes involved in these tasks, rather than to enter into the debate about memory systems as opposed to memory processes. At a minimum, positive transfer, or a dependent relationship between the tasks, was expected.

EXPERIMENT 1

Method

The picture-naming norms of Snodgrass and Vanderwart (1980) were consulted in order to find words that could also be represented as pictures. Five-, six-, and seven-letter words were chosen on the basis of the criterion of 80% or greater agreement on their pictures' names. To obtain norms, fragments and anagrams for these words were shown with a slide projector for 15 and 20 sec, respectively. In order to allow for priming effects to appear, only words with less than 60% completion rates on both of the tasks were selected for use in the experiment. The mean proportion correct was .30 for anagrams and .22 for fragments. This norming procedure also served to eliminate items that had more than one solution.

Four priming or encoding tasks were employed: (1) solving anagrams (e.g., CHANRO), (2) completing word fragments (e.g., ＿＿CHO＿), (3) judging the pleasantness of words, and (4) naming pictures. The retrieval or transfer tasks included (1) solving anagrams, (2) completing word fragments, (3) recognition memory, and (4) immediate free recall. Priming tasks were a between-subjects factor, and transfer tasks a within-subjects factor.

Forty-eight Emory students participated for credit toward a methodology requirement. The subjects were randomly divided into four groups of 12. Group 1 tried to solve 39 anagrams, which were presented on slides for 20 sec each. Group 2 was given 15 sec to solve each of the 39 fragments. Group 3 was presented a list of 39 words for 5 sec each and asked to judge the pleasantness of each one on a five-point scale. Group 4 saw 39 pictures for 5 sec each and was asked to write down the name of each one. The 39 words were the same for all conditions. The subjects wrote each answer on a separate page in a provided booklet. In all but the pleasantness condition, the experimenter read the name of each item after removing the item from the screen. The subjects were told to turn the page in the booklet before each item was named. The subjects did not receive specific instructions to remember the material presented, nor was mention made of the content of the session on the 2d day. Following the presentation of the final item and a brief delay, the subjects were given a free recall test. Completion of this task marked the end of the first part of the experiment.

Twenty-four hours later, the subjects returned to participate in an unannounced memory test. All subjects were given three tasks, each of which included 13 of the words from the previous day's session. The sets of words were counterbalanced according to anagram and fragment difficulty. The task sequence was also counterbalanced across conditions. The anagram and fragment tasks were repeated in the same fashion, except that the subjects were told that some of the words may have been presented during the session on the previous day. In addition to the 13 old words, 7 new ones were added for the anagram and fragment tasks. Since these new items were not chosen as stringently as the 39 words presented at encoding, they were not used to estimate baseline, or unprimed, performance. The other task was a forced-choice recognition test that consisted of three foils for each of the 13 items. The subjects were asked to choose the word they had encountered in the session on the preceding day. The results of the recognition test will not be reported, because of ceiling effects.

Results

Day 1. The subjects solved .38 of the anagrams and .32 of the fragments presented in the first session. The difference between the means (15.0 and 12.3, respectively) fails to reach statistical significance [$t(22) = 1.32$].

Day 2. Each subject provided two measures of priming, one for anagrams and one for fragments. Table 1 (columns 1 and 2) shows the proportion of anagrams and fragments solved during the second session, according to the task engaged in during the first session. A 4 (Day 1 tasks) × 2 (Day 2 tasks) ANOVA showed that performance on Day 2 varied reliably with the method of presentation on Day 1 [$F(3,44) = 7.39, p < .01$]. Neither the type of test on Day 2 (anagrams or fragments) nor the interaction reached significance. All of the aforementioned analyses are based on raw scores in the Day 2 data. As the statistical tests show, these scores depend on the task at encoding. Though the initial difference between anagram and fragment difficulty was not significant, priming above baseline was calculated to compare the fragments with the anagrams better. Performance on the Day 1 task was taken as a common baseline.[1] For example, the subjects who saw anagrams on the 1st day solved .667 of the anagrams on Day 2. When the baseline score for anagrams (the proportion solved on Day 1, .385) is subtracted from the Day 2 proportion (.667), a priming above baseline score of .282 is obtained. On the other hand, the subjects who saw fragments in both stages have a priming score of .357 (.673−.316).

Table 1 presents the priming scores (see columns 3 and 4). One question to be asked of the data pertains to the question of transfer within and between tasks. When the second task involved anagrams, the amount of priming

from prior exposure to anagrams as opposed to fragments led to slightly greater transfer from anagrams (.28 vs. .26); when the second task consisted of fragment completion, the difference favored transfer from prior experience of fragments (.36 vs. .30). Unfortunately, the grouped scores do not allow for a test of the interaction. On another level, the low priming scores on anagrams and fragments following the naming of pictures or the rating of the words for pleasantness suggests that there is considerably less transfer from non-word-oriented tasks on Day 1 to the word-oriented anagram and fragment tasks on Day 2 than there is from working on anagrams and fragments on Day 1. On the other hand, free recall was facilitated by picture naming and judging word pleasantness on Day 1. The mean number of words recalled following anagram solving, fragment completion, pleasantness rating, and picture naming were .37, .36, .46, and .48, respectively [$F(3,44) = 4.42, p < .01$]. Overall, a picture of specificity of operations appears to emerge (see Roediger & Weldon, 1987).

Evidence for statistical dependence was obtained from standard tests comparing the simple and conditional probabilities in both presentation sequences (anagrams followed by fragments and vice versa). For the anagram-fragment sequence, the conditional probability of solving a fragment, given that it was solved as an anagram, was .73, compared with the simple probability of fragment solution of .62 ($\phi = .20$). For the fragment-anagram sequence, the probability of solving an anagram, given that the item was solved as a fragment, was .77, compared with the simple probability of anagram solution of .64 ($\phi = .19$). Chi-square analyses showed the departure from independence to be significant in each case. The dependence that was found contrasts with evidence for independence between perceptual identification and fragment completion obtained by Witherspoon and Moscovitch (1989). To a transfer-appropriate processing theorist, this finding of dependence suggests that the solving of anagrams and the completion of fragments require some of the same operations. To a memory-systems theorists, these results are consistent with the idea that both tasks reflect the operation of a common system. We suggest, however, that the dependence is due simply to the subjects' having been presented with the solution on Day 1.

EXPERIMENT 2

Experiment 2 was designed to examine the role of test priming (Tulving, et al., 1982; Witherspoon & Moscovitch, in press). Unlike in Experiment 1, in Experiment 2, the subjects were not told the solutions after they had attempted to solve the anagrams and fragments. In addition, baseline performance was assessed more appropriately.

Method

Three priming tasks were employed: (1) solving anagrams, (2) solving word fragments, and (3) judging the pleasantness of words. Memory

Table 1
Old Anagrams and Fragments Solved on Day 2,
and Priming Scores (Experiment 1)

	Day 2 Test			
	Proportions Solved		Priming Scores	
Day 1 Condition	Anagrams	Fragments	Anagrams	Fragments
Anagrams	.67	.62	.28	.30
Fragments	.64	.67	.26	.36
Pleasantness	.48	.51	.10	.19
Pictures	.45	.37	.06	.06

Table 2
Old Anagrams and Fragments Solved in Phase 2,
and Priming above Baseline (Experiment 2)

| | Phase 2 Test | | | |
| | Proportions Solved | | Priming Scores | |
Phase 1 Condition	Anagrams	Fragments	Anagrams	Fragments
Anagrams	.33	.34	−.01	.04
Fragments	.25	.28	−.09	−.02
Pleasantness	.41	.50	.07	.20

was assessed by using two transfer tasks: (1) solving anagrams, and (2) completing word fragments. The priming tasks were a between-subjects factor, and the transfer tasks a within-subjects factor. Thirty-six word problems from Experiment 1 were divided into three sets of 12 of equal difficulty. Two of these sets were presented to each of three experimental groups in the initial phase of the experiment. This meant that each group saw 24 items during Phase 1. Each word served as an anagram or fragment for an equal number of subjects. Sixty-three Emory students participated for credit toward a methodology requirement. The subjects were randomly divided into three groups. In the first phase, the material was presented exactly as it was in Experiment 1, except that the solutions were not given for the anagrams or fragments. Following presentation of the final item and collection of their answers, the subjects were given a filler task in the form of a set of mathematical sequences to solve. Phase 2 began after 10 min of the filler task. The subjects were presented 18 anagrams or 18 fragments to solve first, followed by the other task with task order counterbalanced. Each set of 18 items included 12 items that had been presented in the first phase (old items) and 6 new items. The instructions were the same as in Phase 1, except that the subjects were told that they might have encountered some of the items in the previous portion of the experiment. This design allowed a comparison of new and old items in Phase 2.

Results

In Phase 1, the subjects solved .26 of the anagrams and .19 of the fragments. In Phase 2, the subjects solved .34 of the new anagrams and .30 of the new fragments. In neither case is the difference between tasks statistically significant. Overall, Phase 2 performance was significantly higher than performance in Phase 1 [$t(41) = 3.33$].

Table 2 presents performance on the old items in Phase 2. A 3×2 ANOVA on the old items found that performance in Phase 2 varied reliably with the method of presentation in Phase 1 [$F(2,60) = 8.76$, $p < .01$]. A Newman-Keuls post hoc analysis showed that performance in the pleasantness condition was significantly greater than that in the anagram and fragment conditions. Since all items received repetition priming in the pleasantness condition, as opposed to only the solved items in the other conditions, this outcome is not surprising.

Priming above baseline is also shown in columns 3 and 4 of Table 2. Baseline performance was .34 for anagrams and .30 for fragments. These values of priming above baseline are negligible for tasks with priming by anagrams or fragments, and they contrast markedly with the find-

ings in Experiment 1. The most reasonable account for this difference is that auditory presentation of the solution word by the experimenter in Experiment 1 facilitated the priming found there.

Stochastic independence was addressed in the same fashion as in Experiment 1. The conditional probabilities far exceed the simple probabilities (.73 vs. .34 for the anagram–fragment sequence, with $\phi = .48$, and .54 vs. .25 for the reverse sequence, with $\phi = .35$). Chi-square analyses showed the departure from independence to be significant in each case, suggesting that solving the items in the second phase was highly dependent on initial performance. The measures of independence obtained were compared across experiments. In both cases, the amount of dependence in the second experiment was significantly greater ($z = -2.99, p < .01$ for the anagram–fragment order, and $z = -2.29, p < .05$ for the fragment–anagram order). This stronger dependence in Experiment 2 is also observed in the conditional probabilities shown in Table 3: the conditional probabilities of solving an item, given that it was not solved in the first phase, are all above .50 in Experiment 1; but in Experiment 2, the same conditional probabilities range from .15 to .21. This difference is attributable to the lack of any information about the item unless it was solved in the first phase. It is instructive to compare the Experiment 2 proportions in Table 3 for items failed in the first test with performance on new items. For all conditions, performance on failed items is substantially below the baseline proportions correct of .34 for anagrams and .30 for fragments. This result contrasts with the finding of Witherspoon and Moscovitch (1989), whose tests consisted of perceptual identification and fragment completion. They found a deleterious effect for failed items only when the same tasks were repeated, but not when the tasks differed. We found a deleterious effect in both cases.

DISCUSSION

We restrict our remarks here to the issue of test priming. It is clear from the results of Experiment 2 and those of Witherspoon and Moscovitch (1989, Experiment 1) that evidence for test priming is difficult to obtain, at least with the tasks used in these studies. The evidence one looks for is a facilitation of performance on the second test, particularly for those items not answered correctly on the first test, whether correct performance involves the perceptual identification of a briefly exposed word, as in Witherspoon and Moscovitch's study, or the solution of an anagram or fragment puzzle. Clearly, if an item is produced successfully on the first test, then it has effectively been exposed to the subject by virtue of its production, and this exposure may serve as a source of repetition priming. Therefore, the focus for test priming must be on items not produced earlier by the subject or presented by the experimenter. Witherspoon and Moscovitch found no positive transfer for these previously failed items, and neither did we in Experi-

Table 3
Conditional Probabilities for Anagrams (A) and Fragments (F)

Experiment	A_2/A_1	F_2/F_1	A_2/F_1	F_2/A_1	$A_2/\overline{A_1}$	$F_2/\overline{F_1}$	$A_2/\overline{F_1}$	$F_2/\overline{A_1}$
1	.96	.98	.77	.73	.52	.54	.58	.53
2	.83	.80	.54	.73	.15	.17	.17	.21

ment 2. An obvious impediment to finding positive transfer is item selection; failed items are difficult by definition. Furthermore, the substantial correlation across items found here between anagram and fragment difficulty ($r = .56$) indicates that item difficulty is not task-specific. How to demonstrate transfer based solely on operations or procedures—the kind of molecular analysis that a complete transfer-appropriate processing approach would seem to require—may prove a difficult, if not intractable, problem with the kind of word tasks currently in favor.

REFERENCES

JACOBY, L. L. (1983). Remember the data: Analyzing interactive processes in reading. *Journal of Verbal Learning & Verbal Behavior*, **22**, 485-508.

KOLERS, P. A., & ROEDIGER, H. L. (1984). Procedures of mind. *Journal of Verbal Learning & Verbal Behavior*, **23**, 425-449.

ROEDIGER, H. L., & WELDON, M. S. (1987). Reversing the picture superiority effect. In M. A. McDaniel & M. Pressley (Eds.), *Images and related mnemonic processes: Theories, individual differences, and applications* (pp. 151-174). New York: Springer-Verlag.

SNODGRASS, J. G., & VANDERWART, M. A. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

TULVING, E. (1985). How many memory systems are there? *American Psychologist*, **40**, 385-398.

TULVING, E., SCHACTER, D. L., & STARK, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 335-342.

WITHERSPOON, D., & MOSCOVITCH, M. (1989). Stochastic independence between two implicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 22-30.

NOTE

1. To take Day 1 scores as baseline scores is problematic, because it does not allow for possible nonspecific transfer effects. The data of Experiment 2 suggest that the priming estimates in Table 1 may be inflated by .10.