

Type of false recognition and levels of processing in free recall

JOHN H. MUELLER and MATTHEW MARLER
University of Missouri, Columbia, Missouri 65201

Subjects had a single immediate test on each of six successive free recall lists, under instructions to think of either rhymes or synonyms during input. This was followed by final free recall and a final recognition test, the latter involving homonyms and synonyms of words in the primacy, middle, or recency positions of the free recall lists. The results indicated that there was semantic encoding of recency items and that shallow encodings were more durable than expected.

When subjects are required to recall a set of unrelated items, the probability of immediate recall is a function of the item's position in the study sequence, with the last-presented (recency) items recalled best, followed by the initial (primacy) and then the middle items. However, with delayed recall tests a "negative" recency effect is often observed (e.g., Craik, 1970); that is, recency items are less likely to be recalled than items from earlier positions.

One interpretation of the negative recency effect assumes a different level of information processing for recency items compared to earlier items (Craik & Lockhart, 1972). From this view, "shallow" processing involves physical features of the item, for example, orthographic and phonemic attributes, while "deep" processing involves semantic content, and the durability of event memory varies directly with the depth of processing. If the positive recency effect reflects a strategy whereby subjects learn to quickly emit recency items to maximize recall (cf. Maskarinec & Brown, 1974), then it seems likely that such items are not as deeply processed as other items in the list. Positive recency can result in immediate recall because even shallow processing will suffice for the recall of the last-presented items, but negative recency occurs in final recall because recency items were not processed deeply enough to allow consistent and reliable reproduction.

In the present study, a recognition test assessed depth of encoding by comparing the occurrence of two types of false recognitions. The items in the primacy, middle, and recency portions of the free recall lists were words (e.g., URN) that produced both a homonym (EARN) and synonym (VASE). Instead of these critical stimuli, the recognition list included the homonym and the synonym, along with unrelated new words and some old words from the free recall lists; subjects were required to identify the words that had been used in free recall. To the extent that false recognitions reflect

initial encoding of the overlapping attribute (cf. Kausler, 1974), and assuming that synonymy would be a deep feature, then it would be expected that synonyms of recency items would be misidentified as old less often than synonyms of primacy and middle items, since recency critical stimuli were not processed to deep levels.

The prediction for homonym false recognitions depends on the presumed retention characteristics of a shallow encoding. If the shallow encoding survives the delay, then homonyms of recency items should be more frequently misidentified as old than would synonyms of those items, and more often falsely recognized than homonyms from other list positions, since recency items are processed primarily at a shallow level. However, if shallow traces are rapidly lost, as was initially implied (Craik & Lockhart, 1972), there should be no difference as a function of input position, with homonym false alarms only at chance levels.

We also attempted to manipulate depth of processing through various study strategies. Some subjects were instructed to think of rhymes to the words during study, while others were told to think of words with similar meanings. Thinking of synonyms should encourage deeper processing of recency items, with the results that (1) the negative recency effect in final recall would be attenuated, and (2) the synonym false recognition rate would be higher for end items. Thinking of rhymes should make shallow processing more general throughout the list, so that (1) rather than just a negative recency effect there should be an overall deficit in final recall, and (2) more homonym false recognitions would occur for initial items, if shallow encodings endure.

METHOD

Subjects and Design

The 72 subjects were from introductory psychology classes, participating in fulfillment of course requirements. They were divided equally into three groups differing in the special study instructions given for immediate free recall: think of similar sounding words, think of words with similar meaning, or none.

Materials

There were six nonoverlapping 20-item *free recall* lists. The words in Positions 1-4 (primacy), 9-12 (middle), and 17-20

Requests for reprints should be sent to J. H. Mueller. This research was supported in part by funds from the University of Missouri Research Council.

(recency) of Lists 2, 4, and 6 served as critical stimuli for items in the recognition list. Each of the 36 critical stimuli was a word like ALTER which produced both a synonym (CHANGE) and a homonym (ALTAR). The specific words in these positions were rotated through all three segments.

The 144-item *recognition* list was composed of 36 old words from the primacy, middle, and recency positions of Lists 1, 3, 5; 36 new words that were homonyms of the critical stimuli in Lists 2, 4, 6; 36 new words that were synonyms of the critical stimuli, plus 36 new words that were unrelated to the critical stimuli. These item types were balanced throughout the sequence, with the homonym both preceding and following the synonym half the time, at a lag ranging from 10 to 25 items. Items related to different input positions were also distributed throughout the sequence. Two orders of the list were used.

Procedure

Each subject had one trial on each of the six free recall lists. The lists were presented by slide projector at a 2-sec rate, with a single immediate 60-sec written test each time. Following general instructions emphasizing recall in any order, specially instructed subjects were told that thinking of either similar sounding words or words with similar meanings would facilitate recall; they were briefly reminded of these strategies just prior to Lists 2-6.

Each subject was then given a series of forward digit-span tests, requiring about 3 min. Half of the subjects in each group then had a 6-min final free recall test for all 120 items followed by the recognition test, while the task order was reversed for the other subjects. The recognition test used the single-item format; words were presented by slide projector at a 5-sec rate, with the subjects marking their answers on a separate sheet.

RESULTS¹

Immediate Recall

Table 1 summarizes the recall results. The data were analyzed in a 3 (instructions) by 3 (free recall form) by 2 (recognition form) by 2 (task order) by 6 (list) by 3 (position) factorial design, the last two factors being repeated measures. Unless noted otherwise, lists and task order yielded no significant effects.

A marginal instructions main effect revealed that the control group recalled more than the synonyms group, which recalled more than the rhyming group [$F(2,36) = 3.08, p < .06, MSe = 1.16$], and there was the expected main effect due to input position [$F(2,72) = 170.96, MSe = 1.13$]. There was an Instructions by Position interaction [$F(4,72) = 9.29, MSe = 1.13$]; the rhyming strategy depressed primacy recall and increased recency recall, while the synonyms strategy reduced primacy recall slightly without affecting recency, relative to the control condition. The List by Position interaction indi-

cated the expected enhancement of the recency effect with practice [$F(10,360) = 11.84, MSe = .84$]. There was no Instructions by List by Position interaction [$F = 1.06$].

Final Free Recall

The instructions main effect was significant [$F(2,36) = 5.02, MSe = .77$], with the control condition superior to the synonyms group, which was superior to the rhyming group. There was a position main effect [$F(2,72) = 11.53, MSe = .64$], as the primacy items were recalled better than the recency and middle items, which did not differ. There was a marginal Instructions by Position interaction [$F(4,72) = 2.18, p < .08, MSe = .64$], as all groups showed positive primacy, but the rhyming group also showed a positive recency effect. Subjects who had final recall after recognition recalled more than with the reverse order [$F(1,36) = 13.05, MSe = .77$], but there were no interactions with order.

Final Recall as a Function of Immediate Recall

It has been argued that negative recency is more apparent when only items recalled on the immediate tests are scored on the final test (e.g., Engle, 1974; Lockhart, 1975), though Craik, Gardiner, and Watkins (1970) found negative recency even for unrecalled items. The results of this breakdown for the present data are shown in Table 1.

The position main effect indicated negative recency for both recalled and unrecalled items [$F_s(2,72) = 8.51$ and $7.56, MSe = .16$ and $.02$]. The recalled items showed a main effect of instructions [$F(2,36) = 3.53, MSe = .15$], with the control group better than synonym and rhyming strategies. The Instructions by Position interaction was significant for recalled items [$F(4,72) = 2.58, MSe = .16$], with only the control and synonyms groups showing negative recency and positive primacy; this interaction was not significant for unrecalled items [$F = 1.11, MSe = .02$].

Priority in Recall

The priority of items in recall as a function of input position was examined by computing the relative index of priority (Flores & Brown, 1974). The immediate recall data revealed a position main effect [$F(2,72) = 128.97, MSe = .35$], and a List by Position interaction [$F(10,360) = 3.74, MSe = .21$]. Recency items were re-

Table 1
Mean Immediate Free Recall (IFR), Mean Final Free Recall (FFR), Probability of Final Recall Given Immediate Recall (FFR/IFR), and Probability of Final Recall Given Not Recalled on Immediate Test (FFR/IFR*), by Strategy and Input Position

Group	IFR Data			FFR Data			FFR/IFR Data			FFR/IFR* Data		
	Input Position			Input Position			Input Position			Input Position		
	1-4	9-12	17-20	1-4	9-12	17-20	1-4	9-12	17-20	1-4	9-12	17-20
Control	1.99	1.02	2.28	1.15	.83	.84	.46	.39	.29	.06	.07	.06
Rhymes	1.24	1.01	2.53	.79	.64	.83	.32	.29	.33	.07	.06	.02
Synonyms	1.47	1.11	2.30	1.00	.72	.72	.42	.31	.25	.06	.06	.03

called first, followed by primacy and then middle items, with this becoming more pronounced over lists. There were no interactions involving instructions. Final recall indicated no position effect or List by Position interaction ($F_s < 1$).

False Positives

The false positive rates for the homonym and synonym lures are summarized in Figure 1, pooled over task order and lists. Neither the instructions nor position main effects were significant, nor was their interaction ($F_s < 1.79$). Synonym lures led to more false alarms (FAs) than homonyms [$F(1,36) = 3.48, p < .07, MSe = .09$; means = .29 and .26]. There was an Instruction by Type interaction [$F(2,36) = 5.35, MSe = .09$], as the control group had more synonym FAs than homonyms (means = .36 and .26), while the rhyming group (means = .27 and .29) and the synonym group (means = .25 and .24) showed little difference between synonyms and homonyms. The Position by Type and Instruction by Position by Type interactions were not significant ($F_s < 1$).

Control-word FAs revealed a main effect of instructions [$F(2,36) = 4.35, MSe = .02$], with mean rates of .28, .21, and .18 for the control, rhyming, and synonym groups. Special instructions may have helped because, in addition to the familiarity judgment, subjects could also detect that they had not generated rhymes (synonyms) for the unrelated new words.

False Positives by Type

Another analysis separated the number of false positives which involved only the synonym, only the homonym, or cases where both the synonym and homonym were falsely recognized. There was a main effect due to instructions for synonym alone, homonym alone, and both [$F(2,36) = 2.73, p < .08, MSe = .10$; $F(2,36) = 7.19, MSe = .05$; $F(2,36) = 3.02, p < .07,$

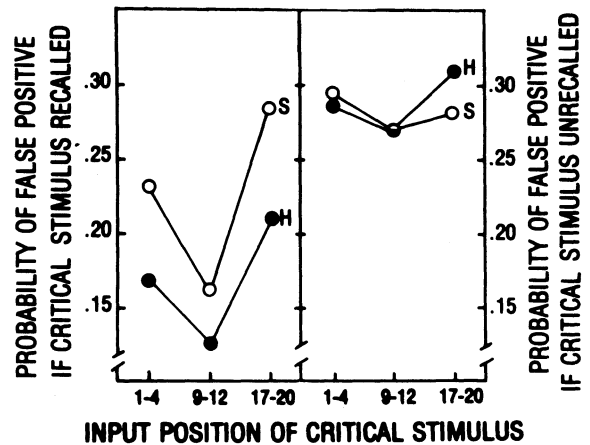


Figure 2. Probability of a false positive as a function of recall of the critical stimulus, by input position of the critical stimulus and distractor type (H = homonym, S = synonym).

$MSe = .05$]. The control group made more synonym-alone and synonym-plus-homonym false positives than the other two groups, while the rhyming group made the most homonym-alone false positives. There were no main effects due to input position ($F_s < 1$). The synonym-alone analysis did indicate an Instructions by Position interaction [$F(4,72) = 2.66, MSe = .02$] that was not present in the other measures ($F_s < 1$). The rhyming strategy depressed synonym-alone FAs to recency items (means = .24, .17, and .20 for primacy, middle, and recency) relative to the control group (means = .25, .24, and .25), while the synonym strategy decreased primacy FAs (means = .14, .19, and .20).

False Positives as a Function of Recalling the Critical Stimulus

False recognitions were also conditionalized upon initial recall by dividing false positives related to items unrecalled on the immediate tests by the number of unrecalled items, and false positives to recalled items by the number of recalled items. The results are shown in Figure 2.

For recalled critical stimuli, the control and rhyming strategy led to somewhat more FAs overall than the synonym strategy [$F(2,36) = 3.12, p < .06, MSe = .17$]. There was a position main effect [$F(2,72) = 10.99, MSe = .11$], as items related to recency critical stimuli led to the most FAs, followed by primacy and then middle items. Neither the Instruction by Position nor the Position by Type interactions were significant ($F_s < 1$). There was a main effect for type of lure [$F(1,36) = 9.22, MSe = .11$], and the Instruction by Type interaction was significant [$F(2,36) = 4.07, MSe = .11$]; synonym lures led to higher FA rates than homonyms in the control group (means = .28 and .15) and with the rhyming strategy (means = .23 and .19), but not with synonyms strategy (means = .15 and .15).

The unrecalled data revealed no position or instruction main effects or interaction ($F_s < 1.94$). The Instruction by Type interaction was significant [$F(2,36) = 4.85,$

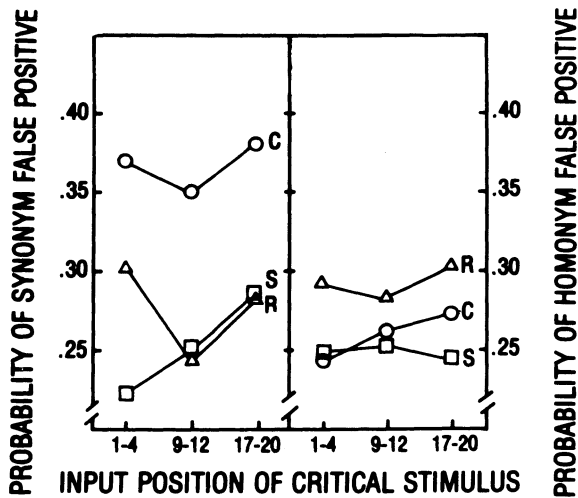


Figure 1. Probability of synonym and homonym false positives as a function of input position of the critical stimulus, for each instructional condition (C = control, R = rhymes, S = synonyms).

MSe = .10]; synonym lures led to more false positives than homonyms for the control group (means = .36 and .30), but the opposite was the case with the rhyming strategy (means = .22 and .30), and there was no difference with the synonyms strategy (means = .26 and .27).

Hit Rates

The hit-rate analysis revealed no position or instruction main effects or interaction ($F_s < 2.11$), with means of .78, .77, and .79, for the primacy, middle, and recency segments.

DISCUSSION

The basic recall data were consistent with previous reports in that the immediate tests revealed the progressive development of recency and priority effects (Maskarinec & Brown, 1974), and the final test revealed negative recency (Craik, 1970). Furthermore, the basic pattern of recognition results was as expected; for example, there were more false positives overall to lures related to primacy items, consistent with several analyses of free recall, including the depth analysis. However, not all of the predictions of the depth analysis were borne out. In particular, these results suggest greater permanence for shallow encodings than was originally assumed (Craik & Lockhart, 1972) and considerable semantic processing of recency items. Evidence converging on the same conclusions has been presented previously (Mueller & Marler, Note 1).

In immediate recall, the rhyming strategy reduced primacy recall and increased recency recall, but it led to the least tendency toward negative recency in final recall. Further evidence for persistence of the shallow encoding was apparent in the recognition data; for example, there were more false positives to homonym lures than to synonyms with the rhyming strategy.

Instructions to think of synonyms during study produced less clear-cut changes. This strategy did reduce immediate primacy recall somewhat, which may reflect the fact that this strategy precluded generalized semantic encoding. That is, generating synonyms may involve deep processing, but it is likely more limited than the broader semantic encoding normally done by control subjects. Otherwise, this group was fairly similar to the control; in particular, there was no reduction of negative recency, nor did false positives to synonym lures increase.

Though the recognition test was delayed, there was little evidence for negative recency; that is, recency critical stimuli led to at least as many false positives as middle items. In fact, when false positives were classified in terms of initial recall or nonrecall of the critical stimulus, recalled items led to more false positives. One explanation for this can be labeled the "reinspection" hypothesis. Briefly, subjects recall recency items both with a high probability and early in output. With written recall, if time remains at the end of the test period subjects may reinspect that record, perhaps using the recalled items as retrieval cues to access unrecalled items. Such scanning accords further study to the recency items, resulting in a more stable memory trace. While admittedly ad hoc here, this analysis does lead to other expectations. For one thing, it suggests that negative recency ought to be greater with oral recall than with written tests, since the subject has no record to reinspect in the former case. Craik (1970) did not specifically report this difference, but reanalysis of the end items in his data revealed more pronounced negative recency with oral recall than with written.²

In the strictest sense, then, only unrecalled critical stimuli conform to the original conception of shallow encoding during input. However, with all due caution in view of the limited number of these items, even unrecalled recency critical stimuli failed to show negative recency (Figure 2). Thus, there still appeared to be a rather permanent memory trace for shallowly encoded (recency) material.

The levels-of-processing view has continued to evolve (e.g.,

Craik, 1977; Craik & Tulving, 1975), and newer versions might incorporate the present results. For example, recent versions argue that retention depends upon the "richness" or elaboration of the encoding *within* a level of processing, as well as depth per se, and this elaboration will be more effective to the extent that the encoding context is *congruous* with the to-be-remembered item. Reduced negative recency in final recall with the rhyming strategy (Table 1) might be explained by assuming that this condition achieved especially rich shallow encodings for recency items. The reduced immediate primacy effect with rhyming (Table 1) might be explained in terms of its incompatibility for such items. The rhyming group's high false alarm rate for recency homonyms (Figure 1) could also be seen as consistent with the idea of an elaborated shallow trace. It seems that problems remain though, such as the substantial semantic processing for recency items (Figure 2). There are clearly limits to how much can be gained by speculations of this sort, and it seems more appropriate to await research specifically intended to test newer versions of the levels-of-processing framework for memory research.

REFERENCE NOTE

1. Mueller, J. H., & Marler, M. *Type of false recognition and levels of processing in free recall*. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1976.

REFERENCES

- CRAIK, F. I. M. The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 143-148.
- CRAIK, F. I. M. Depth of processing in recall and recognition. In S. Dornic & P. M. A. Rabbitt (Eds.), *Attention and Performance VI*. Potomac, Md: Lawrence Erlbaum, 1977.
- CRAIK, F. I. M., GARDINER, J. M., & WATKINS, M. J. Further evidence for a negative recency effect in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 554-560.
- 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., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, 104, 268-294.
- INGLE, R. W. Negative recency in delayed recognition. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 209-216.
- FLORES, L. M., JR., & BROWN, S. C. Comparison of measures of output order in free recall. *Behavior Research Methods & Instrumentation*, 1974, 6, 385-388.
- KAUSLER, D. H. Continuity of processes over variants of recognition learning. In R. L. Solso (Ed.), *Theories of cognitive psychology: The Loyola symposium*. Potomac, Md: Lawrence Erlbaum, 1974.
- LOCKHART, R. S. The facilitation of recognition by recall. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 253-258.
- MASKARINEC, A. S., & BROWN, S. C. Positive and negative recency effects in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 328-334.

NOTES

1. Effects described as significant involve $p < .05$, unless noted otherwise.

2. The authors wish to thank F. I. M. Craik for making these data available for further analysis.