ILLiad TN: 2523

Borrower: TXA

Lending String: *AFU,IAU,IEA,ISM,ITD

Patron: Smith, Steven

Journal Title: Journal of verbal learning and verbal behavior.

Volume: 16 **Issue:**
Month/Year: 1977 **Pages:** 339-352

Article Author:

Article Title: Glenberg, A.M., Smith, S.M., & Green, C; Type I rehearsal; Maintenance and more.

Charge
Maxcost: $50lfm

IL: 3717003

Call #: BF455.A1 J

OCLC#: 1783224

Location:

Shipping Address:
TEXAS A&M UNIVERSITY
ATTENTION: ILL
LOCATION: 41
HUB: HOU

Fax: (979) 862-4759
Ariel: 165.91.220.14

This has been provided b
Libraries at the University
Have a great day!

Date: 2/
Initials:

Per:
Shelf: ✓
Sort:
ILL:

JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **16**, 339–352 (1977)

Type I Rehearsal: Maintenance and More

ARTHUR GLENBERG, STEVEN M. SMITH, AND CAROL GREEN

University of Wisconsin

A technique that can be used to study the effects of low-level, rote, repetitive (Type I) rehearsal is introduced and validated. The technique is then used to investigate the relationship between the amount of Type I rehearsal and recognition memory performance, free from confoundings that have plagued other research. Increasing the amount of Type I rehearsal improved recognition memory regardless of the number of words maintained at one time (one or three) and independent of the association value of CVC syllables. These results imply that Type I processing serves more than a maintenance-only function, and that the additional benefit from this rehearsal is best explained by an increase in the number of frequency or context tags associated with the memorial representation of the to-be-remembered items.

A number of investigations have led to the need to postulate more than one rehearsal process. Within the levels-of-processing framework (Craik & Lockhart, 1972), Type II rehearsal results in deep analysis or mnemonic elaboration. Type II rehearsal is usually characterized as a process which results in the storage of both low-level, structural, and acoustic features as well as semantic features of the stimulus. The amount of Type II rehearsal is positively correlated with delayed memory performance. Craik and Watkins (1973) also propose a maintenance process, Type I rehearsal. This rehearsal process is assumed to maintain information by rote repetition at a given level of analysis for the duration of the processing activity. The amount of Type I rehearsal, however, should not affect the memory trace or delayed memory performance. Working within a multistore framework of memory, a similar distinction has been made by Woodward, Bjork, and Jongeward (1973). They proposed that secondary or associative processing (analogous to Type II processing) results in association formation between to-be-remembered items, and, in effect, transfers information from Short-Term Store to Long-Term Store. They demonstrated that performance on a delayed test improved with the amount of secondary processing. Woodward et al.'s primary rehearsal (analogous to Type I processing) appears to maintain information in Short-Term Store, but does not affect delayed recall performance (see also Craik & Watkins, 1973). Woodward et al. also adduced evidence that primary rehearsal effects some transfer of information to Long-Term Store, specifically, information that can be detected by a recognition test.

This article is concerned with two questions: First, is there any low-level, rote rehearsal process that does not affect recall: and second, given the existence of such a process, what is the effect of this process on memory performance as measured by a recognition test? In order to facilitate the following discussion,

This research was supported by Grant No. 1-R01-MH26643 from NIMH to Arthur Glenberg. We wish to thank Greg Oden for generously taking the time to help us prepare the tapes used to control the stimulus events in these experiments. We also wish to thank Marian Schwartz and Robert Bjork for their comments on an earlier version of this paper. Requests for reprints should be sent to Arthur Glenberg, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

ISSN 0022-537

12

we would like to make a number of distinctions. First, we are going to use the term Type I rehearsal as the generic name for low-level, rote, repetitive rehearsal. Included in this definition are Craik and Lockhart's maintenance rehearsal and Woodward et al.'s primary rehearsal. We will use maintenance rehearsal to refer to a possible process that maintains information for immediate recall but has no effect on delayed memory performance. Primary rehearsal will refer to a possible type of processing that maintains items for immediate recall and shows no effect on a delayed recall test, but the results of this type of processing can be detected with a recognition test. We will use the term Type II rehearsal as the generic name for a different type of processing that encompasses both Craik and Lockhart's second type of rehearsal and Woodward et al.'s associative rehearsal. We will use semantic processing to refer to a possible type of processing that results in the formation of a trace containing semantic features that is not associated with any other traces. Associative rehearsal will be used to denote a type of process that results in associations among memory traces.

Our first question is concerned with the existence of Type I processing. An affirmative answer has been given by Craik and Watkins (1973), Jacoby and Bartz (1972), Modigliani and Seamon (1974), and Woodward et al. (1973). On the other hand, Dark and Loftus (1976), Darley and Glass (1975), Mechanic (1964), and Nelson (1977) have found negative results. There does not seem to be any easy resolution of this disagreement since there have been differences in the paradigms and procedures, definitions of Type I rehearsal, and various artifacts that may have affected the results. Perhaps most difficult is that to prove the existence of Type I rehearsal requires the acceptance of the null hypothesis; the amount of Type I rehearsal should not correlate with recall performance. Demonstrating that the amount of a given type of rehearsal *is* correlated with recall does not

rule out the possibility th
exists that exhibits the chara
I rehearsal. The problem
finding the process.

We propose three criteria
that is to be used to try to el
sal. First, the paradigm m
control of the subject's proce
must be sure that the subjec
maintain the information f
that the subject is not try
associations among the t
units, nor is he trying to inc
analysis during the processi
ious incidental tasks seen
requirement. Perhaps the bes
is one in which the subject
performance is determined
task other than (what is act
sing of to-be-tested inform
subject views as an interferi
type of paradigm the subjec
devote the minimum amou
the processing of the to-be-te
as specified by the instructio
by the task.

The second criterion is th
result in the maintenance o
throughout the duration of t
which the subject is proce
mation. In other words, the
actively holding the inform
with the rehearsal process,
process must be operative
interval. A good paradigm
direct observation of item r
haps through an overt rehea
(Note that we are not, how
that overt rehearsal and Ty
equated.)

The third criterion is tha
Type I rehearsal does not affe
This last requirement introd
into the operational defin
rehearsal. For our purpos
circularity is not totally dam
ment 1 we intoduce a parad

we would like to make a number of distinctions. First, we are going to use the term Type I rehearsal as the generic name for low-level, rote, repetitive rehearsal. Included in this definition are Craik and Lockhart's maintenance rehearsal and Woodward et al.'s primary rehearsal. We will use maintenance rehearsal to refer to a possible process that maintains information for immediate recall but has no effect on delayed memory performance. Primary rehearsal will refer to a possible type of processing that maintains items for immediate recall and shows no effect on a delayed recall test, but the results of this type of processing can be detected with a recognition test. We will use the term Type II rehearsal as the generic name for a different type of processing that encompasses both Craik and Lockhart's second type of rehearsal and Woodward et al.'s associative rehearsal. We will use semantic processing to refer to a possible type of processing that results in the formation of a trace containing semantic features that is not associated with any other traces. Associative rehearsal will be used to denote a type of process that results in associations among memory traces.

Our first question is concerned with the existence of Type I processing. An affirmative answer has been given by Craik and Watkins (1973), Jacoby and Bartz (1972), Modigliani and Seamon (1974), and Woodward et al. (1973). On the other hand, Dark and Loftus (1976), Darley and Glass (1975), Mechanic (1964), and Nelson (1977) have found negative results. There does not seem to be any easy resolution of this disagreement since there have been differences in the paradigms and procedures, definitions of Type I rehearsal, and various artifacts that may have affected the results. Perhaps most difficult is that to prove the existence of Type I rehearsal requires the acceptance of the null hypothesis; the amount of Type I rehearsal should not correlate with recall performance. Demonstrating that the amount of a given type of rehearsal *is* correlated with recall does not rule out the possibility that *some* process exists that exhibits the characteristics of Type I rehearsal. The problem becomes one of finding the process.

We propose three criteria for any paradigm that is to be used to try to elicit Type I rehearsal. First, the paradigm must allow for the control of the subject's processing activity. We must be sure that the subject is trying to only maintain the information for brief periods, that the subject is not trying to form any associations among the to-be-remembered units, nor is he trying to increase the depth of analysis during the processing interval. Various incidental tasks seem to satisfy this requirement. Perhaps the best sort of paradigm is one in which the subject believes that his performance is determined by a concurrent task other than (what is actually) the processing of to-be-tested information, which the subject views as an interfering task. In this type of paradigm the subject is motivated to devote the minimum amount of capacity to the processing of the to-be-tested information as specified by the instructions and demanded by the task.

The second criterion is that the task must result in the maintenance of the information throughout the duration of the interval during which the subject is processing the information. In other words, the subject must be actively holding the information in memory with the rehearsal process, and the rehearsal process must be operative throughout the interval. A good paradigm will allow for the direct observation of item maintenance, perhaps through an overt rehearsal requirement. (Note that we are not, however, suggesting that overt rehearsal and Type I rehearsal be equated.)

The third criterion is that the amount of Type I rehearsal does not affect delayed recall. This last requirement introduces a circularity into the operational definition of Type I rehearsal. For our purposes, however, the circularity is not totally damaging. In Experiment 1 we intoduce a paradigm that satisfies these three criteria for Type I rehearsal. We then use this same paradigm to investigate the effects of this rehearsal process on recognition performance in Experiments 2 and 3.

In our experiments the subjects are led to believe that they are participating in an experiment using the short-term distractor paradigm (Peterson & Peterson, 1959). As far as the subjects are concerned, they are trying to remember four digits. The "distractor" task is the overt repetition of one or three words for 2, 6, or 18 seconds. Following this interval the subjects are required to attempt the recall of the four *digits*. After a number of these digit recall trials the subjects are asked to try to recall the words they rehearsed. The paradigm allows for the manipulation of a number of independent variables including the length of the rehearsal period, the speed of rehearsal (we pace the subject's overt rehearsals), the number of different items rehearsed on each trial, the type of item, and so forth.

This paradigm satisfies our three criteria. Since the subjects do not know that they are to remember the words, there is no reason to believe that they are using any more than the minimal amount of capacity in processing the words. In addition, the subjects are motivated to try to remember the digits, insuring minimal processing of the words. The fact that the items are maintained throughout the rehearsal interval is evident by the overt rehearsal requirement (which also allows a check on misperceptions) without confounding the results of the final recall test with immediate recall tests (see Modigliani, 1976, and Nelson, 1977). Third, as we will demonstrate, recall does not increase with the amount of rehearsal in this paradigm. Finally, the paradigm provides an independent check on the fact that the subjects are doing a minimal amount of processing of the words over and above that specified by the instructions and task demands. If, as we suppose, the subjects are devoting little effort to the processing of the words, then recall of the digits (which the subjects are trying to remember) should be relatively high and show little or no decline over the rehearsal–retention interval. Clearly, this check is not foolproof. If the subject's total capacity is overloaded (either by increasing the number of digits or words), digit performance will suffer. Nonetheless, given a reasonable processing load, performance on the digits should not be affected by the length of the interval.

Our second concern is with the effect of this low-level, rote, rehearsal process, the amount of which does not influence recall, on memory performance as measured by a recognition task. Does the process leave no mark on memory, or is the effect of the process one that cannot be detected by a recall test? Should Type I rehearsal be classified as only a maintenance process (Craik & Lockhart, 1972), or as a primary process that affects recognition memory performance but not recall performance (Woodward et al., 1973)?

Woodward et al. explain the result that Type I rehearsal affects recognition but not recall performance with the help of a two factor theory of recall (e.g., Kintsch, 1970). According to this type of theory, recall involves both retrieval (by some sort of item to item or superordinate to item association) and then a decision (recognition) that the retrieved information was part of the to-be-remembered information. The bottleneck in a recall test is the retrieval stage; recognition of retrieved information is much easier. Since Type I processing does not result in association formation, information studied using Type I rehearsal is difficult to retrieve and recall regardless of the amount of such rehearsal. On the other hand, Type I rehearsal may strengthen the representation of the item in Long-Term Store, or increase the number of frequency or context tags associated with the item's representation. These effects of Type I rehearsal could be detected with a recognition test.

It should be noted that Woodward et al.'s demonstration that Type I rehearsal affects recognition memory performance is at variance with the levels-of-processing specification of Type I rehearsal. According to Craik and Lockhart (1972), the amount of Type I rehearsal should not affect delayed recall or recognition.

Before pursuing any theoretical treatment we believe that the basic finding needs closer examination. In Woodward et al. (Experiment 3), the subjects were shown lists of 36 words. Following each word was a blank interval of 0–12 seconds. At the end of this interval the subjects were required to recall the single word, guaranteeing that the word was maintained for the duration of the interval. After this within-list recall, the subjects were cued to either forget the word or to try to remember the word. At the end of the list the subjects were asked to recall all of the words that were cued to be remembered. The length of the processing interval (0–12 seconds) did not affect this recall. After four lists of 36 words each, the subjects were asked to recall all of the words they could remember. Again, there was no effect of the rehearsal interval on recall. Following this final recall the subjects were given a recognition test. The results from this test indicated that increasing the item-to-cue interval from 0 to 12 seconds produced increases in recognition performance. The design of Woodward et al.'s experiment prevented the subject from knowing whether a word was to be cued to be forgotten or remembered until after the processing interval. Therefore, Woodward et al. reasoned that the subjects would devote very little processing to the word, over and above maintenance (Type I Rehearsal), to avoid interference generated by memorizing to-be-forgotten words. Combining this inference with the results of the final tests leads to the conclusions that the amount of Type I rehearsal does not affect recall but does affect recognition memory performance.

These conclusions are open to a number of criticisms. First, recognition was measured after three different recall tests. While we can think of no specific hypotheses, the recall tests may have confounded the recognition test (Modigliani, 1976; Nelson, 1977). Second, subjects may have used some form of (Craik & Lockhart Type II) semantic rehearsal during the item-to-cue interval. Woodward et al. argue against this criticism since recall did not appear to be affected by the length of the interval. Nonetheless, within the confines of a two-factor theory of recall, their results are compatible with the possibility that the subjects used the item-to-cue interval to continue perceptual or semantic processing of the items. Items cued to be remembered could then have been given some sort of organizational processing that allowed their retrieval and recall. Items that were cued to be forgotten would also have received increased semantic processing during the item-to-cue interval but no organizational processing, hence they were not recalled. In both cases, recognition could have increased with the length of the item-to-cue interval but not recall. Finally, we believe that the most important problem is that the subjects were engaged in a rather complicated task that did require the recall of some of the items. The subjects were probably rehearsing items that were previously cued to be remembered during subsequent item-to-cue intervals. It may be that subjects were trying to use some form of Type I rehearsal, and that Type I rehearsal does not affect recognition memory performance. Nonetheless, the requirement to process more than one item at a time may have forced some semantic (i.e., Craik & Lockhart's Type II) processing during the item-to-cue interval that affected the recognition test and not the recall test.

The basic question is, in a situation devoid of Type II processing, does Type I processing affect recognition memory as suggested by Woodward et al., or does Type I processing only maintain items with no effect on delayed performance, as suggested by Craik and Lockhart? We propose to test four hypotheses that account for the effect of Type I rehearsal on recognition memory. Each hypothesis is framed within a two-factor theory of recall. That is, recall will only increase with processing time when associations or organizations are the product of the process involved. The first two hypotheses state that Type I rehearsal is only a maintenance process; any effect of Type I rehearsal on recognition or recall is due to various artifacts. The remaining hypotheses maintain that Type I rehearsal is a primary rehearsal process and does influence recognition memory.

The Associative-Artifact hypothesis states that Type I rehearsal (maintenance) does not normally improve recognition memory, but when subjects are required to maintain more than one item at a time they cannot avoid some association formation (i.e., Woodward et al.'s associative rehearsal). An episodic association may be formed whenever two items are brought into temporal contiguity, regardless of the subjects' processing activity. This hypothesis, in conjunction with a two-factor theory of recall, predicts that nominal Type I rehearsal will produce increased recognition and recall when subjects must maintain more than one item.

The Depth-Artifact hypothesis also states that Type I rehearsal (maintenance) does not improve recognition memory. When subjects must maintain more than one item, however, they cannot avoid some deeper processing of the items (i.e., Craik & Lockhart's semantic rehearsal), although associations are not formed. The deeper processing may be an unavoidable consequence of maintaining multiple items. On the other hand, the subjects may be forced by the task demands to increase the depth of the representation. A deep representation may be necessary to maintain an item during the interval in which the other items are being rehearsed. Incorporated into a two-factor theory of recall, this hypothesis predicts that nominal Type I rehearsal will produce increased recognition performance when subjects are required to maintain more than one item, but recall will be unaffected by this manipulation.[1] The hypothesis is tested in Experiments 1 and 2.

The Tagging hypothesis assumes that Type I rehearsal (primary) always results in more than the maintenance of information. Specifically, the hypothesis predicts recognition performance to improve with the amount of Type I rehearsal regardless of the number of words rehearsed. The improvement is attributed to the accumulation of frequency or context tags to the memorial representation. Since no associations are formed, this hypothesis predicts no improvement in recall.

The Depth hypothesis also assumes that Type I rehearsal (primary) always results in improved recognition memory regardless of the number of words rehearsed. The improvement is attributed, in this case, to a greater semantic analysis of the information, even though the subjects are engaged in Type I rehearsal. This hypothesis implies that the differences between Type I and Type II rehearsal are not as great as suggested by Craik and Lockhart. Both rehearsal processes may influence the depth of analysis, but Type II processing may also involve association formation (as suggested by Woodward et al.) and quicker processing to a given depth. In conjunction with a two-factor theory of recall, no improvement in recall is predicted by this hypothesis (see footnote 1). These last two hypotheses are tested in Experiments 1, 2, and 3.

[1] A reviewer has commented that it does not make sense to imply that recall will be unaffected by depth of processing in the absence of retrieval cues. Nonetheless, Moscovitch and Craik (1976) demonstrated that the difference in recall between shallow and deep processing increases as the amount of retrieval information is increased. They conclude that "overall retention must be viewed as a joint function of both trace information and retrieval information" (p. 455). We believe that the retrieval information is extremely impoverished when recall is tested in our paradigm, and hence, depth of processing may show no effect. On the other hand, retrieval information (the item itself) is provided on a recognition test, thereby, perhaps, allowing the depth of processing to reveal itself.

If either the Tagging or Depth hypothesis is supported, the usefulness of the theoretically defined maintenance-only Type I rehearsal is called into question. Both of the hypotheses postulate, over and above item maintenance, mechanisms that influence the memorial representation of the information. In addition, support for the Depth hypothesis would question the existence of any process that can maintain information without also increasing the depth of analysis.

Experiment 1

Using the paradigm described above, the subjects overtly rehearsed either one or three words. According to the Associative-Artifact hypothesis, recall in the three-word condition should increase with the amount of rehearsal. The other hypotheses predict no increase. All four hypotheses predict no increase in recall in the one-word condition.

Method

A total of 36 subjects participated in the experiment. Half the subjects were drawn from the Introductory Psychology subject pool at the University of Wisconsin-Madison and received course credit for participation in the experiment. The other subjects were students at the University who were paid $2.50 for their services. All subjects received a bonus dependent on the number of words recalled. The subjects were not informed of this bonus until immediately before the recall of the words. Half of the subjects were assigned to the one-word condition (one word rehearsed in the rehearsal–retention interval) and the other half of the subjects were assigned to the three-word condition. The subjects were run individually.

The stimuli were projected on a wall in front of the subject via a Kodak Carousel projector controlled by one channel of a stereo tape deck. Each trial began with a ready signal of asterisks for 2 seconds followed by a four-digit number for 3 seconds. The next slide presented either one or three words for 2 seconds and was followed by a blank slide. The blank slide remained for 2, 6, or 18 seconds. During this interval the subjects heard a series of 1000-hertz tones, three tones every 2 seconds, that was recorded on the other track of the tape deck. These tones were used to control the frequency of the subjects' overt rehearsals. In the one-word condition, the subject said the word 3, 9, and 27 times in the 2- , 6- , and 18-second rehearsal intervals, respectively. In the three-word condition the subjects rehearsed the set of words 1, 3, or 9 times in the 2- , 6- , and 18-second rehearsal intervals, respectively. In the three-word condition the subjects were instructed to cycle through the words, in order, in time with the tones. Following the rehearsal interval the subjects were allowed 5 seconds to recall the digits after which the ready signal for the next trial was presented.

In the experiment the subjects were exposed to 63 trials. The first three trials were practice trials, one at each rehearsal interval. The next three trials were used to absorb any primacy effects, and the last three trials were used to absorb any recency effects. These nine trials were eliminated from all analyses. The 54 trials in the middle of the sequence were composed of 18 of each each rehearsal interval randomized in blocks of six trials, two of each rehearsal interval.

The words were drawn from a list of 249 single syllable, four- and five-letter common nouns. Sixty-three of the words were drawn for the one-word condition. These 63 words were also used in the three-word condition along with 126 additional words. The 63 words common to both conditions occurred equally often as the top, middle, and bottom words in the three-word condition. In constructing the word triplets for the three-word condition care was taken to insure that the words were not obviously related. In both conditions counterbalancing insured tha each word was used equally often in each rehearsal interval. The remaining 60 words were used as distractors in the recognition test of Experiment 2.

The subjects were instructed to try to remember the digits, in order. They were told that the overt rehearsal task was to prevent them from rehearsing the digits. They were also asked to say one word each time they heard a tone. Following the three practice trials the subjects were allowed to ask questions, and any deviations from the instructions were corrected. The next 60 trials were run with no breaks except to change slide trays. After the last trial subjects were engaged in conversation for a few minutes, and then they were asked to recall the words in any order.

Results and Discussion

On the 54 trials of interest subjects in the one-word condition misperceived or mispronounced only two words out of a possible 972. In the three-word condition the subjects misperceived or mispronounced 54 words. These trials were eliminated from all further data analyses.

Digit recall. Recall was scored as an error if any of the four digits were not recalled or if any were recalled in the wrong serial order. In the one-word condition the mean errors were 7, 7, and 11% in the 2-, 6-, and 18-second rehearsal intervals, respectively. In the three-word condition the mean percentages of error were 24, 26, and 29. The arcsines of the square-root of the error proportions were subjected to a 2 (number of words rehearsed) by 3 (counterbalancing of words to rehearsal intervals) by 3 (rehearsal intervals) analysis of variance, with the last factor being within subjects. The only significant effect was for the number of words rehearsed, $F(1, 30) = 17.96$, $MS_e = 367.08$, $p < .01$. Although the number of errors appears to show a slight increase with the length of the rehearsal interval, it is probably no more than would be expected if the interval was completely empty and the subjects were allowed free rehearsal of the digits. We conclude, therefore, that the subjects were devoting very little effort to the processing of the words.

Since recall of the digits was relatively constant across the rehearsal intervals, the difference between the one- and three-word conditions would seem to have a simple explanation: The difference appears to be due to the forgetting of the digits during the initial encoding of the words, with virtually no forgetting thereafter. Since it is more difficult to encode three words than one word, more of the digits are forgotten in the three-word condition. Interestingly, the number of errors in the three-word condition is just about equal to three times the number of errors in the one-word condition.

Word recall. In the one-word condition subjects' recall averaged 11, 7, and 12% in the 2-, 6-, and 18-second rehearsal intervals, respectively. Considering only the 54 words presented in both the one- and three-word conditions, subjects in the three-word condition averaged recalls of 3, 4, and 5% in the 2-, 6-, and 18-second rehearsal intervals, respectively. The arcsines of the square root of the proportions recalled were subjected to a 2 (conditions) by 3 (counterbalancing of words to rehearsal intervals) by 3 (rehearsal intervals) analysis of variance, with the last factor being within subjects. The only significant source of variability was due to the difference between the one-word and three-word conditions, $F(1, 30) = 11.27$, $MS_e = 130.72$, $p < .01$.

A second analysis was conducted on the recalls in the three-word condition. For this analysis all three words in a trial were scored so that the percentages recalled are based on a total of 54 words per subject (18 trials at each rehearsal interval by 3 words in each trial). As before, the transformed proportions were analyzed. The subjects' average recalls were 7, 5, and 6% in the 2-, 6-, and 18-second rehearsal intervals, respectively. The only significant effect was that for counter-balancing, $F(2, 15) = 3.8$, $MS_e = 44.5$, $p < .05$.

These results are consonant with those

reported by Craik and Watkins (1973) and Woodward et al. (1973). In all of these experiments subjects were induced to maintain items using Type I rehearsal; in all three experiments recall was constant regardless of the amount of Type I rehearsal. In the present experiment the amount of rehearsal devoted to each item increased ninefold between the 2- and 18-second rehearsal intervals, but recall increased a statistically insignificant 1.5%. Just as amazing is the very low absolute level of recall, about 7%, even though the words were given an average of almost nine rehearsals.

Since recall in the three-word condition did not increase with the amount of rehearsal we can reject the Associative-Artifact hypothesis. Increasing the amount of rehearsal should have increased the number of unavoidable associations and improved recall. If these associations were formed they did not appear to increase in frequency over the rehearsal interval nor were they evident in the output order. Considering the 119 words recalled by the subjects in the three-word condition, there were only nine instances of words being recalled (not necessarily contiguous) from the same input triplet. Four of these instances were due to a single subject.

Six possible problems may cloud the interpretation of the results of experiments such as this one. We believe that these problems do not pertain to the present research. The first possible problem is an item selection artifact since we eliminated 54 words from the three-word condition. When these words are included in the data analysis, however, recall of the critical word remains low and constant; 3, 5, and 5% in the 2-, 6-, and 18-second rehearsal intervals, respectively.

The other problems are those listed by Nelson (1977). The first of these concerns the lack of statistical power when accepting the null hypothesis: Is there no improvement in recall, or was our test insensitive to the improvement? In our first analysis, analyzing recall from both conditions, our test had a 95% chance of detecting a true difference between the population means as small as 6%. The average increase between the 2- and 18-second rehearsal intervals was only 1.5%. In the analysis of all three words in the three-word condition our test had a 95% chance of detecting a true difference between the population means as small as 4%. Although the null hypothesis may not be correct, it is very unlikely that recall increases, over a ninefold increase in Type I processing, by more than a few percent.

The second problem Nelson lists stems from floor effects. Two pieces of evidence argu against this criticism in Experiment 1. Firs there was a highly significant differen between the one-word and three-word cond tions. Therefore, even if recall in the thre word condition is so low as to mask a rehearsal interval effects, recall in the o word condition is clearly off the floor. A differences in recall among the rehear intervals should have been evident. T second piece of evidence comes from partiti ing the subjects by their overall performan The three subjects in each counterbalanc that had the highest total recall were selec For the nine best subjects in the one-w condition recall was 15, 11, and 17% in 2- , 6- , and 18-second rehearsal inte conditions respectively. Recall is off floor, and yet there is no improvement with amount of Type I processing.

Nelson's third problem derives from vestigators manipulating the amount rehearsal time, as opposed to the ac number of rehearsals. In those situations investigator has no direct knowledge tha subjects are processing the inform throughout the interval. In the pr experiment the amount of time and the n ber of rehearsals were directly manipul The subjects were overtly repeating the v for the whole interval, and the numb repetitions was held constant across sub

Nelson's fourth criticism concerns co ing, on a final recall, items that were previ called and items that were previously not called. This confounding is absent in Experiment 1, as well as any confoundings introduced y having any previous recall tests of the nformation requested on the final recall.

The final criticism is the lack of replicability of the null finding. In the present experiment we can consider the between-subject condition as providing a replication. Recall was unaffected by the amount of Type I processing for both the one-word subjects and the three-word subjects. In addition, we have been able to replicate the basic result in other experiments conducted in our laboratory.

As noted in the introduction, there appear to be three criteria for Type I rehearsal. Our paradigm has met these criteria while avoiding common confoundings. We conclude, therefore, that a Type I rehearsal process does exist, and hence, we proceeded to investigate the effect of Type I rehearsal on recognition memory performance.

Experiment 2

We included both the one-word and three-word conditions in this experiment. The results of Experiment 1 did not support the notion that association formation occurred when more than one word was maintained. Nevertheless, the requirements to maintain more than one word may force some Type II rehearsal in the sense of greater semantic analysis without concurrent association formation, to which a recall test is not sensitive (Depth-Artifact hypothesis). If this hypothesis is correct, and if the probability of a correct recognition depends on the depth of processing, then recognition performance in the three-word condition should increase, while performance in the one-word condition remains constant. The Tagging and Depth hypotheses predict increases in both the one-word and three-word conditions.

Method

The 63 trials were exactly the same as in Experiment 1. At the end of the last trial the subjects were asked to try to recognize the words they had rehearsed. Sixty words, one from each trial excluding the practice trials, along with 60 new words, were presented to the subjects on sheets containing 12 to 16 words. Next to each word was a scale from 1 to 6 which the subjects used to mark their recognition judgment (from positive it is old to positive it is new). The first 12 words on the test represented one word from each of the six buffer trials mixed with six distractors. The subjects were allowed unlimited time to complete the test; however, they were instructed to judge one word at a time and not to go back and change their judgments.

The 36 subjects were drawn from students in the introductory psychology classes at the University of Wisconsin-Madison. None of the subjects had served in Experiment 1. The subjects were given course credit for their participation in the experiment. Half of the subjects were assigned to the one-word condition, and the other half to the three-word condition. The subjects were run individually.

Results and Discussion

On the 54 trials of interest, subjects in the one-word condition misperceived or mispronounced a total of two words. In the three-word condition the subjects missed 62 words. These trials were eliminated from all further data analyses.

Digit recall. Recall of the digits was very similar to that in Experiment 1. In the one-word condition subjects erred on 6, 6, and 7% of the trials in the 2- , 6- , and 18-second rehearsal intervals, respectively. In the three-word condition the errors averaged 24, 29, and 31% in the 2- , 6- , and 18-second rehearsal intervals, respectively. The proportions were analyzed as in Experiment 1, and again the only significant source of variability was that for the number of words rehearsed, $F(1, 30) = 27.16$, $MS_e = 340.81$, $p < .01$. Again, it seems reasonable to conclude from these data that the subjects were spending little effort on the words, over and above maintenance.

Word recognition. The mean hit rates, d's (measure of discrimination from signal detection theory, Green & Swets, 1966) and confidence ratings are presented in Table 1. The individual d's were subjected to a 2 (number of words rehearsed) by 3 (counterbalancing) by 3 (length of the rehearsal interval) analysis of variance, with the last factor being within subjects. The difference between the one-word and three-word condition was significant, $F(1, 30) = 32.53$, $MS_e = .91$, $p < .01$. The effect of the rehearsal interval was also significant, $F(2, 60) = 5.28$, MS_e .10, $p < .01$. The length of the rehearsal interval did not interact with the number of words being rehearsed, $F(2, 60) = .41$. The only other significant effect was the number of words rehearsed by counterbalancing interaction, $F(2, 30) = 3.63$, $p < .05$. Although performance in the one-word condition was always superior to performance in the three-word condition, the differences between the average d's ranged from .39 to 1.60 across the three counterbalancings. Similar analyses were conducted on the hit rates and confidence ratings that yielded the same pattern of results.

In this paradigm, which has satisfied our three criteria for Type I processing, recognition performance increases with the amount of rehearsal. Contrary to the Depth-Artifact hypothesis, the three-word condition did not have a greater slope than the one-word condition. In fact, the slope is slightly less than in the one-word condition. The same result is obtained when comparing equal number-of-rehearsal conditions. In the three-word condition, individual words were rehearsed three and nine times in the 6- and 18-second rehearsal intervals, respectively, while d's increased by .12. In the one-word condition individual words were rehearsed three and nine times in the 2- and 6-second rehearsal intervals, respectively, while d′ increased by .14. From these findings we conclude that Type I rehearsal results in more than just maintenance.

A possible problem with this conclusion is that the subjects in the three-word condition experienced a change in context between learning and testing (Tulving & Thomson, 1973). During learning the words were rehearsed in triples; at the test only a single word from the triple was judged as old or new. There are a number of reasons for discounting this problem. First, the finding that the words in the three-word condition were recognized more poorly than the words in the one-word condition seems very reasonable given the direction of the difference in Experiment 1. Second, there is no reason to believe that a context change would differentially affect the slopes of the recognition performance functions in the two conditions even though the absolute performance levels may be affected. Finally, the experiment was designed to test the Depth-Artifact hypothesis against the

TABLE 1

MEAN HIT RATES, d′s, AND CONFIDENCE RATINGS IN EXPERIMENT 2

Performance measure	One-word condition Rehearsal interval (seconds) 2	6	18	Three-word condition Rehearsal interval (seconds) 2	6	18
Hit rate	.65	.67	.74	.56	.59	.63
d′	1.35	1.49	1.67	.38	.44	.67
Confidence rating	4.23	4.34	4.67	3.76	3.88	4.04

Note: The average false alarm rates equaled .19 in the one-word condition and .42 in the three-word condition.

Tagging and Depth hypothesis. The Depth-Artifact hypothesis not only predicts an interaction between the number of words rehearsed and rehearsal time, it also specifies that performance in the one-word condition will not be affected by the amount of rehearsal. The Depth-Artifact hypothesis may be eliminated by finding no interaction (as above) or by rejecting the null hypothesis that rehearsal time does not affect performance in the one-word condition. An analysis of the d's from the one-word condition indicates a significant main effect of rehearsal time, $F(2, 30) = 3.49$, $MS_e = .13, p < .05$.

The question now is why recognition improves with the amount of Type I rehearsal. We had suspected that forcing subjects to maintain two or more words in memory concurrently would lead to a type of processing that would improve recognition, while Type I processing of a single word would have no effect. The results of Experiments 1 and 2 contradict this hypothesis. The two other hypotheses seem to provide reasonable explanations of the improvement, and they are tested in Experiment 3.

Experiment 3

If the increase in recognition performance is due to the addition of frequency tags with Type I rehearsal (the Tagging hypothesis), then the increase in performance should be found regardless of the type of material rehearsed. Specifically, if CVC syllables are rehearsed, then the improvement in recognition with increased Type I rehearsal should not interact with the meaningfulness of the CVCs.

On the other hand, Craik and Lockhart (1972) suggest that (for Type II processing) meaningful stimuli "will be processed more rapidly than less meaningful stimuli" (p. 676). If Type I processing affects the depth of processing (the Depth hypothesis), and if greater depth is achieved more rapidly with meaningful stimuli than with less meaningful stimuli, then rehearsal time should interact with the meaningfulness of the stimuli. The form of the interaction, however, cannot be accurately specified. If the high meaningful stimuli are processed to an asymptotic depth very quickly (e.g., within 2 seconds), then performance on these stimuli would not appear to depend on the length of the rehearsal interval (2 to 18 seconds). Assuming a slower processing rate for the low meaningful stimuli, performance would depend on the length of the rehearsal interval. If, however, the asymptotic depth of analysis of the high meaningful stimuli is not reached quickly (e.g., not until 6 or more seconds of processing), then performance on both the high and low meaningful stimuli will increase with the length of the rehearsal interval, with performance on the high meaningful stimuli increasing fastest.

In sum, any interaction between meaningfulness and processing time would contradict the Tagging hypothesis and lend support to the Depth hypothesis. The form of the interaction, if found, will provide information on the rates of processing for these materials. The lack of an interaction will provide weak evidence contrary to the Depth hypothesis and in favor of the Tagging hypothesis.

Method

In this experiment the subjects were treated exactly the same as in the one-word condition of the previous two experiments except that CVC syllables were used instead of words. A set of 123 nonword CVC syllables was drawn from the pronunciability norms in Hayden and Loud (1969) and Underwood and Schulz (1960). All the stimuli had a pronunciability rating under 4.0 on a nine-point scale, with a rating of 1 being easiest to pronounce. The syllables (excluding the three syllables used in the practice trials) were divided into two sets of 60 high association value syllables (Glaze, 1928, values of 70–100), and 60 low association value syllables (0–67). The syllables in the two sets had approximately the same distribution of first letter frequencies. Each set of 60 items was then subdivided into two sets of 30 items. One set of 30 items was used as the distractors; the items in the other set were presented during the rehearsal intervals. Again, the frequency of first letter usage was approximately the same in the presentation sets and the distractor sets. The mean pronunciability and association values for the four sets of items are as follows: high association value presented items, 2.47, 86.52; high association value distractors, 2.43, 87.59; low association value presented items, 2.79, 42.44; low association value distractors, 2.48, 41.0.

Each subject contributed nine observations to each of the six conditions formed by crossing rehearsal interval (2, 6, or 18-seconds) and association value (high or low). Counterbalancing insured that within each association value the same CVC occurred equally often at each rehearsal interval. In addition, each of the 60 trials was used equally often for low and high association value syllables.

The subjects were 36 students at the University of Wisconsin-Madison who had not served in the previous experiments. The subjects were paid $2.50 for their services. As in previous experiments, the subjects were run individually.

Results and Discussion

Digit recall. On the trials in which a low association value syllable was presented, the subjects erred on 11, 10, and 14% of the digits in the 2- , 6- , and 18-second rel vals, respectively. When high value syllables were presentec errors were 9, 11, and 13% in the 18-second rehearsal intervals, The arcsines of the square root of tions were analyzed in a 6 (counte $\times$ 2 (association values) $\times$ 3 (reh vals) analysis of variance, with i factors being within subjects. significant effect was the triple $F(10, 60) = 2.05$, $MS_e = 88.64$, p interaction is due to unsystematic v digit recall over the counterbalanci theless, it appears to be safe to co the form of the syllable rehearsed a length of the rehearsal interval effect on digit recall.

Syllable recognition. The mean h and confidence ratings are present 2. The d's were calculated using the rates to the appropriate associa distractors. The analysis of varianc was the same as that used on the d The only two significant sources of were association value, $F(1, 30) = 9$ $.59, p < .01$, and rehearsal time, $F(2,$ $MS_e = .25$, $p < .025$. The interactic association value and rehearsal tin significant, $F(2, 60) = 1.66$, $MS_e =$ In fact, the sums of squares for action (.70) is so small compared t mean square that no single contrast

TABLE 2

MEAN HIT RATES, d's, AND CONFIDENCE RATINGS IN EXPERIMENT 3

Performance measure	High association value Rehearsal interval (seconds)			Low association val Rehearsal interval (	
	2	6	18	2	6
Hit rate	.68	.72	.74	.62	.61
d'	1.13	1.22	1.29	.80	.76
Confidence rating	4.38	4.53	4.62	4.08	4.03

Note: The average false alarm rate equaled .28 for the high association value syllables and .32 association value syllables.

was then subdivided into two sets of 30 items. One set of 30 items was used as the distractors; the items in the other set were presented during the rehearsal intervals. Again, the frequency of first letter usage was approximately the same in the presentation sets and the distractor sets. The mean pronunciability and association values for the four sets of items are as follows: high association value presented items, 2.47, 86.52; high association value distractors, 2.43, 87.59; low association value presented items, 2.79, 42.44; low association value distractors, 2.48, 41.0.

Each subject contributed nine observations to each of the six conditions formed by crossing rehearsal interval (2, 6, or 18-seconds) and association value (high or low). Counterbalancing insured that within each association value the same CVC occurred equally often at each rehearsal interval. In addition, each of the 60 trials was used equally often for low and high association value syllables.

The subjects were 36 students at the University of Wisconsin-Madison who had not served in the previous experiments. The subjects were paid $2.50 for their services. As in previous experiments, the subjects were run individually.

Results and Discussion

Digit recall. On the trials in which a low association value syllable was presented, the subjects erred on 11, 10, and 14% of the digits in the 2-, 6-, and 18-second rehearsal intervals, respectively. When high association value syllables were presented, the digit errors were 9, 11, and 13% in the 2-, 6-, and 18-second rehearsal intervals, respectively. The arcsines of the square root of the proportions were analyzed in a 6 (counterbalancings) × 2 (association values) × 3 (rehearsal intervals) analysis of variance, with the last two factors being within subjects. The only significant effect was the triple interaction, $F(10, 60) = 2.05$, $MS_e = 88.64$, $p < .05$. The interaction is due to unsystematic variations in digit recall over the counterbalancings. Nevertheless, it appears to be safe to conclude that the form of the syllable rehearsed and that the length of the rehearsal interval had little effect on digit recall.

Syllable recognition. The mean hit rates, d's, and confidence ratings are presented in Table 2. The d's were calculated using the false alarm rates to the appropriate association value distractors. The analysis of variance on the d's was the same as that used on the digit errors. The only two significant sources of variability were association value, $F(1, 30) = 9.69$, $MS_e = .59$, $p < .01$, and rehearsal time, $F(2, 60) = 4.61$, $MS_e = .25$, $p < .025$. The interaction between association value and rehearsal time was not significant, $F(2, 60) = 1.66$, $MS_e = .21$, $p > .1$. In fact, the sums of squares for the interaction (.70) is so small compared to the error mean square that no single contrast can result

TABLE 2

MEAN HIT RATES, d's, AND CONFIDENCE RATINGS IN EXPERIMENT 3

Performance measure	High association value Rehearsal interval (seconds)			Low association value Rehearsal interval (seconds)		
	2	6	18	2	6	18
Hit rate	.68	.72	.74	.62	.61	.73
d′	1.13	1.22	1.29	.80	.76	1.10
Confidence rating	4.38	4.53	4.62	4.08	4.03	4.46

Note: The average false alarm rate equaled .28 for the high association value syllables and .32 for the low association value syllables.