Memory & Cognition
1980, Vol. 8 (6), 528-538

Spacing repetitions over 1 week

ARTHUR M. GLENBERG and THOMAS S. LEHMANN
University of Wisconsin, Madison, Wisconsin 53706

Does varying the spacing of repetitions over intervals as long as 1 week affect recall? The answer from three experiments is yes. Subjects incidentally processed words repeated within a single list and words repeated in separate lists at list spacings of up to 1 week. Memory was tested by free recall shortly after the second presentations or after retention intervals of up to 1 week. Recall of the words repeated across separate lists conformed to a proportionality rule. When the retention interval is short relative to the spacing intervals, performance is inversely related to spacing. When the retention interval is a large proportion of the spacing intervals, performance is directly related to spacing. Does varying the spacing of repetitions within a single list affect recall after a retention interval of 2 weeks? The answer depends on the processing used while studying the words. Processing that generated interitem associations resulted in a within-list spacing effect even after a 2-week retention interval. Without the interitem associations, the effect was absent after a 1-day retention interval. Most of these findings were explained by examining the changing relationship between the retrieval context and the context stored during study.

As a general rule, as the spacing between the presentations of repeated information is increased, memory performance also increases. This phenomenon is called the spacing or lag effect. Although there has been considerable theoretical and empirical research on the spacing effect (for reviews, see Hintzman, 1974; Melton, 1970), using a variety of procedures, most of the interrepetition intervals investigated have been relatively modest. The spacings of the repetitions found in the vast majority of the studies range from 1 or 2 sec to a few minutes. The one exception of which we are aware is a study reported by Landauer and Ross (1977). They asked subjects to memorize a seven-digit telephone number over the course of a 1-week interval by using their own methods (control group) or by briefly studying the number once or twice daily for 1 week (spaced practice group). Two weeks after receiving these instructions the control group had 61% perfect recalls and the spaced practice group had 73% perfect recalls of the number.

The use of long spacing and test intervals has two important benefits. First, it allows us to look at the generality of the spacing effect under conditions approaching those of everyday life (at least in the time intervals, if not in the stimulus material). Second, it provides a fresh series of tests for theoretical explanations of the spacing effect. In the present series of experiments, words are repeated after delays of up to 1 week. Memory for the information is tested by free recall after retention intervals of up to 2 weeks.

The experiments reported in this paper were designed to answer two questions. First, will recall after extremely long spacings (spacings of 1 day or 1 week) conform to a proportionality rule that holds for recall after short spacings? The proportionality rule is that when the retention interval is short relative to the spacing of the repetitions, performance is negatively correlated with repetition spacing; when the retention interval is long relative to the spacing intervals, performance is positively correlated with spacings of the repetitions.

Data conforming to the proportionality rule were first reported by Peterson, Hillner, and Saltzman (1962) using the continuous paired associate paradigm. These paired associate effects were replicated and extended by Glenberg (1976). In addition, the proportionality effect has been found in free recall using standard spacings of 0-60 sec (Glenberg, 1977) and in free recall after substantially longer spacings. In Glenberg's (1979) second experiment, words were repeated in different lists with spacings of 3.7 min to 25.9 min. Free recall was requested approximately 1.4 min or 120 min after the second presentation. Consonant with the proportionality rule, after the short retention interval, recall decreased as the spacing of the repetitions increased, whereas after a long retention interval, spacing and recall were directly related. The present experiments are extensions of those reported earlier, with intervals measured in days instead of minutes. A theoretical interpretation of these proportionality effects is presented in the general discussion.

The second question concerns the effects of extremely long retention intervals on typical spacing effects. When the repetition spacing is varied within a standard free

Experiments 1 and 2 are part of the second author's senior honors thesis. Part of those data were presented at the meetings of the Midwestern Psychological Association, May 1979. We thank Anthony Gino for helping us collect and analyze the data in Experiment 3. This research was supported by NIMH Grant MH26643 and University of Wisconsin-Madison Graduate School Research Committee Grant 190244 to Arthur Glenberg. Requests for reprints should be sent to Arthur Glenberg, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

0090-502X/80/060528-11$01.35/0

recall list and retention is tested shortly thereafter, performance is strongly affected by spacing. Given that recall is off the floor, will the within-list spacing of repetitions still influence recall after a retention interval of 1 day, 1 week, or 2 weeks? For theories that attribute the spacing effect solely to enhanced storage of widely spaced repetitions (e.g., Hintzman, 1976; Jacoby, 1978), the answer is affirmative. For theories that attribute the spacing effect to the interaction between storage and retrieval processes (e.g., component-levels theory, Glenberg, 1979), the answer is different.

According to component-levels theory, increasing the spacing between repetitions results in an increasing amount of contextual information stored with the representation of the repeated item. This is referred to as differential storage. Recall is assumed to be a positive function of the match between this stored contextual information and the contextual information available at the test (in the absence of other retrieval cues, such as experimenter-provided cues or subject-generated associative cues based on previously constructed subjective organizations). The context change brought about by increasing the retention interval decreases the match between the test context and the differentially stored input context. Therefore, with sufficiently long retention intervals, the theory predicts that the within-list lag effect will be eliminated.

To test this prediction, these experiments all used incidental learning procedures designed to minimize interitem association formation. Under these conditions, the presence of a within-list lag effect after very long retention intervals is a disconfirmation of component-levels theory and support for storage theories. The absence of such an effect is, conversely, support for component-levels theory and disconfirming evidence for the storage theories.

EXPERIMENT 1

In this experiment, the subjects were presented two lists, either in the same or in different sessions. The lists were composed of word pairs.[1] The subjects were asked to generate a meaningful relation between the words in a pair and then to rate the difficulty of finding the relation. Subjects were not told that their memories for the pairs would be tested. The incidental procedure was used to prevent rehearsal between sessions and to avoid strategies that would result in the storage of associations between the pairs.

Some pairs were repeated within a list either immediately after the first presentation (Lag 0) or after a lag of 5-10 intervening pairs (Lag 7). This within-subjects variable is referred to as within-list lag.

Some pairs were presented in both lists. These lists were separated by a few minutes or 1 week. This between-subjects variable is referred to as between-lists spacing. A second between-subjects variable was the retention interval. The test was given either a few minutes after the second list was presented or after 1 week. Hence four between-subjects conditions were defined. In Condition 0-0 (0-day between-lists spacing and 0-day retention interval after the second list), subjects attended one session in which two lists were presented and performance was tested. In Condition 0-7 (0-day between-lists spacing and 7 days between the second list and the test), subjects attended two sessions separated by 1 week. Both lists were presented in the first session, and performance was measured in the second. In Condition 7-0, subjects also attended two sessions separated by 1 week. A single list was presented in the first session. The second list was presented in the second session, and performance was tested in the second session. In Condition 7-7, subjects attended three sessions with 1 week between the first and second sessions and 1 week between the second and third sessions. The two lists were presented in the first two sessions, and performance measured in the third.

At the test, the subjects were asked to free recall the pairs. According to component-levels theory, the within-list lag effect is expected for the pairs repeated within each list (at Lags 0 and 7), but the size of the effect should diminish as the retention interval between the list and the test increases. Statistically, this prediction translates into a four-way interaction among within-list lag, first or second list, between-lists spacing, and retention interval. The effect of within-list lag should be large for the first and second lists in Condition 0-0 (both lists are tested in the same session in which they are presented) and the second list in Condition 7-0 (which is presented and tested in the same session). The within-list lag effect should be small in the first list in Condition 7-0 and in both lists in Conditions 0-7 and 7-7. All of these lists have substantial (i.e., at least 7-day) retention intervals.

Recall of those items repeated between the lists tests the proportionality rule. If the rule holds for 1-week spacing intervals, then recall in Condition 0-0 should be superior to recall in Condition 7-0 and recall in Condition 0-7 should be inferior to recall in Condition 7-7.

Method

Materials and Design. Two presentation sequences were constructed, each containing 58 serial positions. The first five positions were filled with once-occurring primacy pairs, and the last five positions were filled with once-occurring recency pairs. The 48 middle positions were filled with 8 pairs given two contiguous (Lag 0) presentations, 8 pairs given two presentations separated by 5 to 10 presentations of other pairs (Lag 7), and 16 pairs presented once in the list. These 16 pairs occurred in both lists. The only difference between the two sequences was in the exact assignment of the three types of pairs to serial positions.

The pairs were constructed from a pool of common four- and five-letter single-syllable nouns. Words in a pair were not obviously related. Pairs were randomly assigned to the primacy and recency positions, as well as to the critical repeated positions. The critical pairs were divided into three sets of 16 pairs.

One set was used for the pairs repeated between the lists. Two sets were used for the pairs repeated within the lists, one set for those given Lag 0 repetitions (eight in each list structure) and one set for those given Lag 7 repetitions (eight in each list structure). The three sets of pairs were rotated among the three repetition conditions so that each pair was in each condition for one-third of the subjects. The pairs were printed on punched cards that were sequenced into the lists.

The subjects were assigned to one of the four major conditions, 0-0, 0-7, 7-0, and 7-7. Within each group, half of the subjects were presented the two lists in one order and the remaining subjects viewed the lists in the opposite order. Within each of these subgroups, one-third of the subjects had each assignment of sets of pairs to repetition conditions. This basic design called for 24 subjects (4 groups × 2 list orders × 3 assignments of sets), and we intended to complete the design three times.

Procedure. The subjects were run in groups of 1 to 18. Each subject was handed a booklet including a consent form, three rating forms (only two were used for rating the pairs), and a deck of punched cards. Subjects were told that they were to rate each pair as to how easy the two words were to associate. Specifically, they were to read the pairs, think of a meaningful connection or relation between the words, judge the difficulty of finding the relationship on a scale of 1-5, and then record this judgment on the prepared form. A number of illustrations of meaningful relations were provided using interactive imagery, sentence mediation, and so on. Ten seconds were allowed for the rating of a pair, after which a number, prerecorded on a tape recorder, was heard. The number was the cue to turn to the next punched card (pair). In addition, the number indicated the serial position of the pair, which corresponded to the number printed in the left-hand corner of the card and the number on the rating score form.

Following the instructions, the subjects went through the rating procedure on five cards (which consisted of pairs of two- and three-letter words) for practice. After any questions were answered, the first 58 pairs were presented and rated. Following the last pair, the decks of cards were collected and the subjects in the 7-0 and 7-7 conditions were excused, with the reminder to return in 1 week for a similar task. When these subjects returned, they were treated identically to the 0-0 and 0-7 subjects who remained: The decks for the second list were distributed, and an abbreviated form of the original instructions was read. All subjects rated a second practice list of five pairs before the second list.

After rating the pairs in the second list, 0-7 and 7-7 subjects were excused and reminded to return in 1 week. When these subjects returned, they were treated identically to the 0-0 and 7-0 subjects who remained in the laboratory after the second list. The effective retention interval for Conditions 0-0 and 7-0 was approximately 2 min, the time needed to collect the second set of punched cards and read the recall instructions. All subjects were instructed to use the back of the third rating sheet to free recall the pairs and any individual words they remembered from both lists. Subjects were allowed up to 15 min for recall, after which they were debriefed.

Subjects. Men and women students enrolled in introductory psychology classes at the University of Wisconsin, Madison, participated in the experiment as one way of fulfilling a class requirement. Although the design called for a total of 72 subjects, 18 in each condition, the total number of subjects in each group varied. Subjects were excluded from the various conditions to maintain the counterbalancing scheme. The data reported are from 48 subjects, 12 in each of the conditions.

Results and Discussion

The data were scored and analyzed for complete recall of pairs and for the number of correctly recalled individual words. The two analyses yielded nearly identical patterns of main effects and interactions. Only the data based on individual word scoring are presented. In all statistical tests, the probability of a Type I error is set at .05.

Consider first the recall of the pairs repeated between the lists. The relevant means are presented in the right-hand column of Table 1. According to the proportionality rule, immediate recall of the pairs should be better in Condition 0-0 than in Condition 7-0. Delayed recall of the pairs should be superior in Condition 7-7 compared with Condition 0-7. The interaction of between-lists spacing and retention interval was significant [$F(1,44) = 25.72$, $MSe = 7.47$]. There was a significant between-lists spacing effect at both retention intervals, but in opposite directions. After a 0-day retention interval, recall of items with 0-day between-lists spacings (.42) was significantly greater than recall of items with 7-day between-lists spacings (.24) [$t(22) = 4.06$, $SE = 1.37$]. After a 7-day retention interval, recall after 0-day between-lists spacings (.06) was significantly inferior to recall after 7-day between-lists spacings (.14) [$t(22) = -3.11$, $SE = .78$].

These results provide a (somewhat amazing) extension of the proportionality rule. The rule holds for intervals as short as a few seconds and as long as 1 week. This consistency appears to reflect some basic facts about retrieval. A theoretical interpretation of these effects is provided in the General Discussion section.

Consider now the pairs repeated within a list at Lags 0 and 7. Analysis of the data revealed significant main effects for list spacing [$F(1,44) = 9.76$, $MSe = 4.13$], retention interval [$F(1,44) = 103.07$, $MSe = 4.13$], first or second list [$F(1,44) = 15.26$, $MSe = 5.93$], and within-list lag [$F(1,44) = 18.28$, $MSe = 1.73$].

There were also a number of significant interactions: List Spacing by Retention Interval [$F(1,44) = 22.63$, $MSe = 4.13$], List Spacing by First or Second List [$F(1,44) = 4.55$, $MSe = 5.93$], Retention Interval by Within-List Lag [$F(1,44) = 6.92$, $MSe = 1.73$], Retention Interval by First or Second List [$F(1,44) = 7.11$, $MSe = 5.93$], List Spacing by Retention Interval by Within-List Lag [$F(1,44) = 6.92$, $MSe = 1.73$], and List Spacing by Retention Interval by First or Second List [$F(1,44) = 8.43$, $MSe = 5.93$].

All of the main effects and interactions in the recall of the pairs repeated within the lists were modified by

Table 1
Mean Proportions of Repeated Words Free Recalled in Experiment 1

List 2 Retention Interval*	Repetition Locus: Within Lists, Lag 0	Within Lists, Lag 7	Between Lists, Lag 0	Between Lists, Lag 7	Between Lists
0-0	.22	.40	.27	.35	.42
7-0	.07	.04	.23	.32	.24
0-7	.04	.02	.03	.06	.06
7-7	.06	.10	.06	.05	.14

**In days.*

the significant four-way interaction [$F(1,44) = 9.45$, $MSe = 2.26$]. The form of the interaction was exactly as predicted by component-levels theory. There was a substantial within-list lag effect in both lists in Condition 0-0 and in the second list in Condition 7-0, whereas the effect was absent in all other lists. Unfortunately, recall in these latter lists was so low that the absence of a lag effect may simply reflect a floor effect. Experiment 2 was designed to raise the level of recall to examine this interaction more closely.

EXPERIMENT 2

The design of this experiment was similar to that in Experiment 1, but with a few changes. First, we tried to boost the overall level of recall by using single words instead of pairs. This change necessitated a different orienting task. Subjects were told to generate a superordinate category for each word, write down the category, and then rate the word on an activity dimension. Second, we included in each list once-presented words. If there is no within-list lag effect, these words will prove useful in determining if recall of the twice-presented items reflects a floor effect. They are also useful in assessing the level of recall of the words repeated between the lists in comparison with the level predicted by assuming independent representations. The third change was to include a between-lists spacing of 1 day in addition to the 0-day and 7-day between-lists spacing conditions.

Method

Materials and Design. Two presentation sequences were constructed, each containing 68 serial positions. The first six positions were filled with once-occurring primacy words, and the last six positions were filled with once-occurring recency words. The 56 middle positions were filled with 8 words presented once, 8 words given two contiguous (Lag 0) presentations, 8 words given two presentations separated by 5 to 10 presentations of other pairs (Lag 7) and 16 words presented once in the list and repeated in the other list. The only difference between the two sequences was in the exact assignment of these four types of words to serial position.

The words were drawn from a pool of common four- and five-letter single-syllable nouns. Words were randomly assigned to the primacy and recency positions, as well as to the four types of words presented in the body of the list. These critical words were divided into four sets of 16 words each. One set was used for the words repeated between the lists. Two sets were used for the words repeated within the lists, one set for those given Lag 0 repetitions (eight in each list structure) and one set for those given in Lag 7 repetitions (eight in each list structure). One set was used for the words presented once (eight in each list structure). The four sets were rotated among the four conditions so that each word was in each condition for one-fourth of the subjects. The words were printed on punched cards that were sequenced into the lists.

The subjects were assigned to one of the six major conditions: 0-0, 1-0, 7-0, 0-7, 1-7, and 7-7. Within each group, half of the subjects were presented the two lists in one order and the remaining subjects viewed the lists in the opposite order. Within each of these subgroups, one-fourth of the subjects had each assignment of sets of words to conditions. This basic design called for 48 subjects (6 groups × 2 list orders × 4 assignments of sets), and we completed the design two times.

Procedure. The subjects were run in groups of 4 to 16. Each subject was handed a booklet including a consent form, three of the rating forms (only two were used for the orienting tasks), and a deck of punched cards. Subjects were told that they were to rate each word on an activity dimension by assigning a number between 1 and 7 to the word. The number 1 indicated that the word represented an inactive concept such as "stone" or "death"; 7 indicated that the word represented an active concept such as "explosion" or "football." Following the rating, subjects were asked to generate and write down a superordinate category for the rated word. Associates were allowed if a superordinate could not be generated. Ten seconds were allowed for both tasks, after which a number, prerecorded on a tape recorder, was heard. The number was the cue to turn to the next punched card (word). In addition, the number indicated the serial position of the word, which corresponded to the number printed in the left-hand corner of the card and the number on the response form.

Following the instructions, the subjects went through the rating procedure on five cards (which consisted of three-letter words) for practice. After any questions were answered, the first 68 words were presented and rated. Following the last word, the decks of cards were collected and the subjects in the 1-0, 7-0, 1-7, and 7-7 conditions were excused, with the reminder to return after an appropriate interval for a similar task. When these subjects returned, they were treated identically to the 0-0 and 0-7 subjects who remained: The decks for the second list were distributed, and an abbreviated form of the original instructions was read. All subjects rated a second practice list of five words before the second list.

After rating the words in the second list, the 0-7, 1-7, and 7-7 subjects were excused and reminded to return in 1 week. When these subjects returned, they were treated identically to the 0-0, 1-0, and 7-0 subjects who remained in the laboratory after the second list. The effective retention interval for these conditions was approximately 2 min, the time needed to collect the second set of punched cards and read the recall instructions. All subjects were instructed to use the back of the third rating sheet to free recall the words they remembered from both lists. Subjects were allowed up to 15 min for recall, after which they were debriefed.

Subjects. A total of 96 men and women students enrolled in introductory psychology classes at the University of Wisconsin, Madison, participated in the experiment as one way of fulfilling a class requirement.

Results and Discussion

Recall data for the words repeated between the lists are presented on the right of Table 2. The proportionality rule was again supported by the significant interaction

Table 2
Mean Proportions Free Recalled in Experiment 2

List 2 Retention Interval*	Repetition Locus: Within Lists, List 1: 1P	List 1: Lag 0	List 1: Lag 7	List 2: 1P	List 2: Lag 0	List 2: Lag 7	Between Lists
0-0	.39	.41	.58	.50	.56	.72	.70
1-0	.18	.26	.26	.46	.52	.65	.56
7-0	.08	.19	.27	.53	.57	.68	.54
0-7	.07	.08	.21	.07	.11	.17	.22
1-7	.09	.14	.27	.13	.10	.24	.32
7-7	.09	.07	.13	.11	.11	.12	.25

Note—*1P = pairs presented once.* **In days.*

of between-lists spacing and retention interval [$F(2,48) = 5.87$, $MSe = 6.10$]. The form of the interaction was not, however, exactly as predicted. After a 0-day retention interval, recall decreased as the between-lists spacing increased, as expected. After a 7-day retention interval, recall first increased and then decreased with spacing. The interpretation of this nonmonotonicity is not clear. Also significant in this analysis are the main effects of retention interval [$F(1,48) = 113.59$] and a counterbalancing variable, the assignment of words to conditions [$F(3,48) = 3.15$]. This variable did not interact with any other variable and did not affect any of the conclusions.

There was an interesting relationship between the levels of recall for the words repeated between the lists and the words presented once (1P in Table 2). Assume the words repeated between the lists are given two independent representations in memory, one when presented on List 1 and one when presented on List 2. We might expect the proportion recalled to be given by $P_1 + P_2 - P_1P_2$, where P_1 is the probability of retrieving the representation formed during List 1, and P_2 is the probability of retrieving the representation formed during List 2. The proportion of once-presented words recalled from List 1 provides an estimate of P_1, and the proportion of once-presented words recalled from List 2 provides an estimate of P_2. Using these proportions, the predicted recalls were .70, .56, and .57 for the 0-0, 1-0, and 7-0 conditions, respectively, almost exactly the same as the observed recall. Based on these data, it appears that the items repeated between the lists generate two independent representations. This conclusion changes, however, after considering recall after a 7-day retention interval. Assuming independence, the predicted proportions were .14, .21, and .19 for the 0-7, 1-7, and 7-7 conditions, respectively. In these cases, observed recall was substantially higher than predicted by independence. A statistical test of the difference between predicted and observed proportions was conducted by computing a predicted proportion for individual subjects.[2] The differences between observed and predicted proportions were then used to test the null hypothesis of no difference. The hypothesis could be rejected for the 0-7 and 1-7 conditions [$ts(15) = 2.10$ and 3.01, respectively]. For the 7-7 condition, the null hypothesis could not be rejected [$t(15) = 1.22$].

One way for storage theories to account for this pattern is to assume that the between-lists words start out with two independent representations and then, over the course of the retention interval, they coalesce into a single representation that has a strength greater than either alone. Although this sounds a bit magical, there are some scenarios that could produce it. For example, subjects may inadvertently (or for some perverse reason) recall some of the words during the retention interval. The probability of recalling a word repeated between the lists would be given by the independence equation. Nonetheless, once recalled, the representation of the between-lists word may automatically retrieve its counterpart representing the occurrence on the other list. This processing may, then, produce a link between the representations or coalesce the representations into one likely-to-be-recalled memory.

Another explanation of the switch from independence to nonindependence depends on considering retrieval. Assume that the representations of a word repeated between the lists are not independent, but related. For example, the same cognitive representation may include information encoded from both of the presentations. In this case, the independence in the recall after a 0-day retention interval may indicate independent retrieval attempts using two retrieval cues, not independent storage. After a 7-day retention interval, the retrieval environment may have changed so that only one of the cues is effective, leading to dependence in retrieval. An explanation along these lines is offered in the General Discussion.

The data from the words repeated within each list are also presented in Table 2. Contrary to the prediction from components-levels theory, there was a within-list lag effect in almost every list, regardless of the retention interval. The effect was significant [$F(1,48) = 64.35$, $MSe = .90$], and it did not interact with any variable in a way that changes these conclusions. The predicted four-way interaction was far from significant ($F < 1$).

A number of variables and interactions were significant, although for theoretically trivial reasons. These are listed for completeness: main effects for retention interval [$F(1,48) = 247.78$, $MSe = 2.64$], between-lists spacing [$F(2,48) = 5.95$], and first or second list [$F(1,48) = 99.55$, $MSe = 1.23$], and interactions of retention interval and between-lists spacing [$F(2,48) = 6.58$], retention interval and first or second list [$F(1,48) = 108.94$], between-lists spacing and first or second list [$F(2,48) = 7.61$], and the three-way interaction of these variables [$F(2,48) = 6.88$]. In addition, there were a number of three-way and four-way interactions involving the counterbalancing variables (order of list presentation and assignment of words to conditions). None of these interactions changed the direction of any of the major effects in Table 2; they only indicate relative attenuation and enhancement of those effects.

The consistency of the within-list lag effect indicates that it is likely to be due to enhanced storage of words repeated after long lags, rather than to an interaction between storage of contextual components and their retrieval. This finding can be accommodated within component-levels theory if, in addition to storing contextual information at each presentation, subjects were also storing interitem associations. These interitem associations can provide a source of retrieval cues in addition to the test context. Furthermore, once access to the list is achieved, the use of these previously stored associative retrieval cues should be independent of the test context, resulting in a within-list lag effect at all retention intervals.

This possibility led us to search for evidence of the

storage of interitem associations and their use as associative retrieval cues. First, we noted that when subjects generated superordinates as part of the orienting task, they used some of the superordinates repeatedly. When subjects generated associates (when they could not generate a superordinate), they, at times, generated previous list words. These observations do indicate that although we tried to design the orienting tasks to focus the subjects on the individual stimuli, the tasks may have resulted in a good deal of interitem association. The next experiment was another attempt to avoid a floor effect for recall and simultaneously limit the storage of interitem associations.

EXPERIMENT 3

This experiment was also designed to test the proportionality rule and measure the within-list lag effect after long retention intervals. To maintain recall at reasonable levels, shorter lists were used, and all the stimuli were high-frequency high-imagery words. The orienting task was redesigned to prevent differential (as a function of within-list lag) storage of interitem associations. At each presentation, a word was judged to be more or less costly than a standard (house) or larger or smaller than a standard (dog). These standards were constant throughout the experiment and hence would not be particularly effective retrieval cues. When a word was repeated, it was judged on the same or on a different dimension.

The same-different variable provides a test of an interpretation of the lag effect. Some theorists propose that the effect is due to inattention (Underwood, 1969) or superficial processing (Jacoby, 1978) at the second of two massed presentations. If these are the causes of the lag effect, then having a different task at the second presentation should increase attention and processing (relative to the same task) and reduce the lag effect. Similarly, some encoding variability interpretations of the lag effect (e.g., Maskarinec & Thompson, 1976) predict that requiring different processing at the first and second presentations will produce equivalent encoding variability in the representations of words repeated at all lags. This equivalency should attenuate the lag effect. In component-levels theory, however, changes in the variability on an encoding dimension influence the lag effect only when components from that dimension are an important part of the retrieval cue. In Experiment 3, the retrieval cue was the context at the test. No information relating to the different encodings produced by the judgment task was provided at the test. Therefore, the theory predicts no interaction of repetition spacing and the same-different variable in free recall.

In contrast to the previous experiment, in this experiment, between-lists spacing was a within-subjects variable. Each subject saw three lists. The first and second lists were always separated by 1 day. The third list was presented immediately after the second list on the 2nd day. Some words were presented on both the first and third lists (1-day between-lists spacing). Some words were presented on both the second and third lists (0-day between-lists spacing). The interval between the third list and the recall test was a between-subjects variable. Half of the subjects attempted to recall all of the words immediately after the presentation of the third list. The other subjects attempted recall 1 day after processing the third list.

Method

Materials and Design. Three presentation sequences were constructed, each containing 56 serial positions. The first four positions were filled with once-occurring primacy words, and the last four positions were filled with four once-occurring recency words. The middle 48 positions were composed of four blocks of 12 positions each. Within each block, there were two words each presented once (one with the size judgment and one with the cost judgment), two words repeated within the block at a lag of zero intervening items, two words repeated within the block at lags of two to six intervening items (called Lag 4), and two words whose second presentations were in the third list. In each block, one of each type of repeated word had the same judgment task on its two presentations. The four combinations of first or second presentation and size or cost judgment were used equally often within a list and approximately equally within a block. The only differences between the list structures that were used for the first two lists were in the assignment of conditions to serial positions within a block. The structure of the third list was the same as the other two, except that the serial positions otherwise reserved for once-occurring words and words repeated between the lists were all filled with the second presentations of the words repeated between the lists.

The stimuli were selected from the Paivio, Yuille, and Madigan (1968) norms to have imagery values greater than 4.2 on a 7-point scale, on which 7 indicates the highest degree of imagery. The stimuli also had Thorndike-Lorge (1944) frequency values of at least 26. The standards "house" and "dog" were selected so that about half of the words represented objects more expensive than a house and about half of the words represented objects larger than a dog.

Within Lists 1 and 2, the critical words were divided into eight sets of four words each. Eight assignments of sets of words to conditions were made so that each word was used equally often in one of the eight conditions within a block. In List 3, only four assignments were made.

Subjects were randomly assigned to the two major between-subjects conditions with either a 0-day retention interval or a 1-day retention interval. Within each group, half the subjects were presented the first two lists in one order and the remaining subjects viewed the lists in the opposite order. Within each of these subgroups, one-eighth of the subjects had each assignment of sets of words to presentation conditions.

This basic design called for 32 subjects (2 groups × list orders × 8 assignments of words to conditions); we intended to complete the design once. Analysis of the data after 32 subjects indicated, however, a number of significant crossover interactions involving counterbalancing variables and theoretically significant independent variables. At that point, the words were regrouped into eight new sets of four, and another 32 subjects were tested. The data from these 32 subjects were more orderly. Only trends that replicated across sets of subjects will be reported.

Procedure. The subjects were run in groups of three to eight. Each subject was handed a booklet including a consent form, three rating forms, and a deck of punched cards. Each punched card listed, in order, the serial position of the card, the to-be-judged word, the appropriate judgmental dimension, and the standard for that dimension. Subjects indicated their judgments by writing a plus (larger or more expensive) or minus on the rating forms. A prerecorded audiotape paced the judgments at one every 7 sec.

Following the instructions, the subjects practiced on 12 cards using three-letter words. After any questions were answered, the first list was presented. Following the rating of the last word, all subjects were excused, with the reminder to return to the laboratory in 24 h for a similar session. When these subjects returned, they were read an abbreviated form of the instructions and were presented the second list. The cards for the second list were then collected, and the decks for the third list were distributed. Following the presentation of the third list, subjects who were to have a 1-day retention interval were excused, with the reminder to return the next day. When they returned, they were treated identically to those subjects with the 0-day retention interval, who remained in the laboratory. All subjects were instructed to try to free recall the words from all three lists. Subjects were given up to 15 min for recall, after which they were debriefed.

Subjects. A total of 64 men and women students enrolled in introductory psychology classes at the University of Wisconsin, Madison, participated in the experiment as one way of fulfilling a class requirement.

Results and Discussion

The data of interest are shown in Table 3. For completeness, the data are displayed for words processed using the same orienting task at both presentations, words processed using different orienting tasks at both presentations, and the combined data. Differences between sections of Table 3 are small and unsystematic. Therefore, attention should be focused on the bottom section, containing the combined data.

Again, the data from the words repeated between the lists supported the proportionality rule. The interaction of between-lists spacing and retention interval was significant [$F(1,56) = 8.78$, $MSe = .71$]. The simple main effects were both significant, with the probability of a Type I error set at .06. After a 0-day retention interval, recall of words repeated with 0-day between-lists spacings was greater than recall of words repeated with 1-day spacings [$t(31) = 1.98$, $SE = .33$]. After a 1-day retention interval, recall of words repeated with 0-day spacings was less than recall of words repeated with 1-day spacings [$t(31) = 1.94$, $SE = .31$]. The proportionality effect was constant across replications, and it was unaffected by the same-different variable (both $Fs < 1$). As in the previous experiments, the effect of retention interval was also significant [$F(1,56) = 9.08$, $MSe = .99$].

A second analysis was performed on the words repeated within Lists 1 and 2. Component-levels theory predicts a within-list lag effect for the words repeated in List 2 and recalled in the same session (0-day retention interval) and no, or reduced, effects of lag in the other three lists. The data in Table 3 conform to these predictions, but the interaction of within-list lag, first or second list, and retention interval did not quite reach significance [$F(1,56) = 3.37$, $MSe = .55$, $p = .07$]. A separate analysis of the recall of the words repeated in List 2 and recalled after a 0-day retention interval demonstrated that the lag effect in that list was significant [$F(1,28) = 6.67$, $MSe = .42$]. This within-list lag effect did not interact with the same-different variable ($F < 1$). Apparently, the within-list lag effect was not the result of inattention or deficient processing at the second of two massed presentations.

This pattern of results is consistent with the predictions of component-levels theory, but it is still necessary to demonstrate that the absence of the within-list lag effect in the lists with long retention intervals was not due to floor effects. One index that recall of the repeated words does not reflect a floor effect is provided by demonstrating that recall of the repeated words is statistically superior to recall of the items presented

Table 3
Mean Proportions Free Recalled in Experiment 3

Retention Interval* for Lists 2 and 3	Repetition Locus							
	Within Lists						Between Lists	
	List 1			List 2				
	1P	Lag 0	Lag 4	1P	Lag 0	Lag 4	0 Days	1 Day
				Same Task				
0	.10	.07	.09	.20	.23	.29	.39	.33
1	.11	.20	.16	.12	.18	.14	.23	.38
				Different Task				
0	.10	.15	.06	.20	.21	.30	.44	.34
1	.11	.19	.22	.12	.21	.22	.26	.26
				Combined				
0	.10	.11	.08	.20	.22	.30	.41	.33
1	.11	.19	.19	.12	.20	.18	.24	.32

*Note—1P = words presented once. Because the same-different task classification does not apply to once-presented items, the proportions based on all of the items are reported in the three sections. *In days.*

once. This index fails when applied to recall from List 1 in the 0-day retention interval condition. After a 1-day retention interval, however, average recall of the twice-presented words was significantly greater than the recall of the once-presented words [for List 1, $F(1,56) = 10.13$, $MSe = .951$; for List 2, $F(1,56) = 5.71$, $MSe = 1.055$]. We conclude, therefore, that when interitem associations are eliminated, the within-list lag effect is attenuated after substantial retention intervals and that the attenuation is not an artifact of low levels of recall.

Recall of the words repeated within List 3 presents two mysteries. First, there is no within-list lag effect ($F < 1$). Second, there is a highly significant effect of same-different orienting task [$F(1,56) = 18.64$, $MSe = .75$]. Average recall of words repeated with the same task was .20; for the different tasks, average recall was .32. The only other significant effect in the analysis of List 3 was the retention interval [$F(1,56) = 38.11$, $MSe = .98$], with average recall after a 0-day retention interval equaling .36 and average recall after a 1-day retention interval equaling .17.

GENERAL DISCUSSION

Can spacing repetitions over 1 day or 1 week facilitate performance relative to spacings of a few minutes? The answer from all three experiments is yes. In fact, the long spacings improved performance by 25% to 130% over the "short" spacings. More important than simply illustrating long spacing effects, all three experiments demonstrated the proportionality effect: Performance increases with spacing when the retention interval is long relative to the spacing intervals; performance is inversely related to spacing when the retention interval is short relative to the spacing interval. In free recall, the proportionality effect has been demonstrated using intervals measured in seconds, hours, and now days. These various proportionality effects can be encompassed within a framework provided by component-levels theory (Glenberg, 1979).

The theory proposes that when a word is presented for study, three types of informational components are encoded and that the variability in the encoding of all components increases as a function of repetition spacing. This increase in stored information as a function of spacing is referred to as differential storage. The descriptive components represent the encoding of the word on phonemic, graphemic, semantic, and imaginal dimensions to which the subject is attending. Previous research (Glenberg, 1979) has shown that differential storage of these components has little effect on the magnitude of the spacing effect in free recall. The absence of an interaction between spacing and same-different task in Experiment 3 is consistent with this previous research.

Structural components represent the interitem associations the subject imposes on the presented information. The incidental learning procedures used in Experiments 1-3 were designed to attenuate differential storage of these components by focusing processing on individual items.

The third type of encoded component represents the cognitive context in which the information is encoded. Subjects are somewhat aware of their environment, their feelings, and their extraexperimental thoughts while participating in an experiment. Components representing this awareness (perhaps the contents of short-term store) are assumed to be encoded automatically upon the construction of a memory trace.

The contextual components are constantly changing ("drifting," in Bower's, 1972, terminology, and "fluctuating," according to Estes, 1955); not all aspects of the context change at the same rate, however. The contextual components can be aligned on a dimension defined by the rate of change. Toward one end of the continuum are the local contextual components that represent quickly changing aspects of the context (e.g., the subject's transient thoughts). Because they change quickly, they are included in few traces. Toward the other end of the continuum are the global contextual components. These components encode aspects of the context that change more slowly (e.g., the subject's affective state). Higher order contextual components (more toward the slowly changing end of the continuum) represent perceived aspects of the context that change very slowly, if at all (e.g., the environmental context, Smith, 1979). These components are included in many traces.

As the spacing of presentations increases, there are correlated changes in the context and, consequently, in the number of different contextual components encoded and represented in the trace of the repeated item. That is, there is differential storage of contextual components.

In comparing two spacing conditions, both of which are relatively short, there is differential storage of local contextual components. There is little variability in the storage of global components because there is little change in the global context between the presentations of items repeated after short spacings. In comparing two spacing conditions, both of which are relatively long, there is differential storage of the global components. The two long spacing conditions have equivalent storage of local contextual components; for both conditions, the presentations occur in disparate local contexts, so that different sets of local contextual components are encoded at the presentations.

According to this theory, retrieval begins by encoding or constructing a retrieval cue from the information available at the test. In free recall, this information consists of the encoding of the context at the time of the test. Traces are activated in direct proportion to the number of contextual components shared by the trace and the cue. It is also assumed that the ability of individual components in the cue to activate the stored trace is inversely related to the number of traces includ-

ing that component (i.e., cue overload, Watkins & Watkins, 1975). Thus global contextual components are less effective retrieval cues than are local contextual components.

Traces that are activated beyond a threshold value can be decoded and examined. The information in the trace can be used to support recall or a recognition judgment, as well as being incorporated into the retrieval cue to activate other traces. In this way, structural components (when stored) can be used to access one trace from its associate (a trace that includes the same structural component).

How the theory accounts for the proportionality rule can be illustrated by considering two items, A and B. Item A is presented with a relatively short interval between its first (P_1) and second (P_2) presentations. Item B is presented with a longer P_1-P_2 interval. Reading the following two paragraphs, while assuming that Items A and B are repeated within a free recall list, generates an explanation of within-list proportionality effects (Glenberg, 1977). Note that under these conditions, differential storage of the local context is important, since the items encode the same global contextual components. An explanation of the between-lists proportionality effect (Experiments 1-3) is generated by reading the following two paragraphs and assuming that Items A and B are repeated between the lists and by substituting "global" for "local."

First, suppose that for both A and B the retention interval (P_2-test) is very short compared with the P_1-P_2 intervals. In this case, the local (global) context at the test will share components with the local (global) encoding context at P_1 and P_2 for Item A. For Item B, the local (global) test context will only share components of the encoding context at P_2; P_1 for Item B is encoded in a local (global) context different from the test. The local (global) context at the test will, therefore, activate Item A more than Item B, and recall will be inversely related to the spacing interval.

When the P_2-test interval is long (compared with the P_1-P_2 interval), the local (global) context on the test will have fluctuated or drifted so that it is no longer highly similar to the P_2 encoding contexts. Indeed, because of this fluctuation, the local (global) context available at the test is assumed to include components randomly selected from the local (global) encoding contexts. Now the probability that a component on the test is included among the encoded components of a trace increases directly with the total number (variability) of components stored. Since Item B was presented at a long P_1-P_2 interval, its representation encodes more local (global) contextual components than the representation of Item A, the representation of Item B will be more strongly activated by the test context, and thus the positive spacing effect is produced.

In effect, the whole range of proportionality effects can be explained by applying the same storage and retrieval processes at different points along the rate-of-context-change continuum.

There are two results that are consistent with the context-matching explanation of the proportionality effect. First, in Experiment 2, the words repeated between the lists were recalled at a level predicted by the independence equation when recall was tested after a short retention interval. Following a long retention interval, the level of recall exceeded that predicted by independence. It was suggested that these results reflected the use of two independent retrieval cues at the short retention interval and a shift to different cues at the longer retention interval. These various retrieval cues can be given an interpretation within component-levels theory. After a short retention interval, the effective cue consists of both the local and global contextual components available at the test. After a longer retention interval, the effective retrieval cue consists solely of the global context available at that test. The local contextual components available at the delayed test will not match those encoded at input.[3]

Second, Ross and Landauer (1978) propose a test, or index, of context retrieval that involves once-presented items. They propose a correspondence between a single word presented twice and two words each presented once. Increasing the spacing between the presentations of the repeated word increases the total number of different contextual components stored in that word's representation. Similarly, increasing the spacing between the once-presented words increases the total number of different contextual components stored in the two separate representations. The degree to which the context at the test matches the contextual elements stored in the representation of the repeated word is the same as the degree to which the context at the test matches the contextual elements in one or the other of the once-presented words. As Ross and Landauer note, if context matching produces the spacing effect, then any prediction regarding the effect of spacing that is made for recall of the repeated words must also be made for recall of at least one (one or the other or both) of the two once-presented words. Therefore, the once-presented words should generate the between-lists proportionality effect. These predictions were tested using the data in Experiment 2 from Conditions 0-0, 1-0, 0-7, and 1-7.

Each of the once-presented words in one list was paired with a once-presented word in the other list. We then tallied the number of times a subject recalled one, the other, or both members of the pairs (an OR score) in the four conditions. The average OR scores are .66 and .53 in Conditions 0-0 and 1-0, respectively. The average OR scores are .12 and .20 in Conditions 0-7 and 1-7, respectively. The "spacing effects" for the OR scores are exactly the same as the spacing effects for the twice-presented words (see Table 2). Both follow the proportionality rule. An analysis of variance on the OR scores resulted in a significant interaction of between-lists spacing and retention interval [$F(1,32) = 4.63$, $MSe = 2.28$]. These data, along with the independence analysis, increase our confidence that the between-lists spacing

effects are due to a type of context matching, as described by component-levels theory.

It should be noted that Ross and Landauer (1978) did not find a spacing effect in their OR scores and thus rejected the context-matching explanation of spacing effects. One likely cause of this difference in results is that they examined within-list OR scores, whereas the analysis presented above used between-lists OR scores. Perhaps context matching is an adequate explanation of between-lists spacing, but not of within-list spacing. Alternatively, Ross and Landauer's results may have been due to their subjects' adopting a retrieval strategy other than context matching, such as using previously recalled words as cues to other words.

Do spacings of a few seconds influence performance after retention intervals of days or weeks? The answer seems to depend on the type of information differentially encoded at the presentations and whether or not that information is available at the test. According to component-levels theory, when subjects engage in activities that generate interitem associations (structural components), increasing the spacing of repetitions increases the number of interitem associations stored. Retrieval of some of the words (through global contextual cues) provides access to these structural components, thus generating the positive spacing effect even after long retention intervals. Apparently, this is the cause of the within-list spacing effect found at all retention intervals in Experiment 2.

When the storage of interitem associations is prevented, the theory proposes that the within-list lag effect in free recall depends on the local contextual components in the retrieval cue matching the differentially stored local contextual components. After a long retention interval, however, the local context at the test is unrelated to the stored local contextual components. Since the items repeated within a list do not differentially encode global contextual components, the global context at the test produces equivalent recall of these items. Thus, in Experiment 3 (in which interitem associations were eliminated), the within-list lag effect disappeared when recall was requested after a 1-day retention interval.

More generally, we believe that all of these data provide a strong incentive to closely examine how the test context is used in retrieval and how changes in the context affect what is remembered. The theoretical notions we have discussed in this paper provide an explanation of the proportionality rule in studies of spacing. Similar ideas have also been used to explain the long-term recency effect (Glenberg, Bradley, Stevenson, Kraus, Tkachuk, Gretz, Fish, & Turpin, in press), and they appear to have a wide range of other applications (e.g., Raaijmakers & Shiffrin, Note 1).

REFERENCE NOTE

1. Raaijmakers, G. W., & Shiffrin, R. M. *Search of associative memory.* Unpublished manuscript, 1980.

REFERENCES

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington, D.C: Winston, 1972.

Estes, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, **62**, 145-154.

Glenberg, A. M. Monotonic and non-monotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 1-15.

Glenberg, A. M. Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, **3**, 282-294.

Glenberg, A. M. Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, 1979, **7**, 95-112.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., Fish, J. H., & Turpin, B. A. A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory,* 1980, **6**, 355-369.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola symposium.* Potomac, Md: Erlbaum, 1974.

Hintzman, D. L. Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10). New York: Academic Press, 1976.

Jacoby, L. L. On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 649-667.

Landauer, T. K., & Ross, B. H. Can simple instruction to use spaced practice improve ability to remember a fact? An experimental test using telephone numbers. *Bulletin of the Psychonomic Society*, 1977, **10**, 215-218.

Maskarinec, A. S., & Thompson, C. P. The within-list distributed practice effect: Tests of varied context and varied encoding hypotheses. *Memory & Cognition*, 1976, **4**, 741-747.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 596-606.

Paivio, A., Yuille, T. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology: Monograph Supplement*, 1968, **76**(1, Pt. 2), 1-25.

Peterson, L. R., Hillner, K., & Saltzman, D. Supplementary report: Time between pairings and short-term retention. *Journal of Experimental Psychology*, 1962, **64**, 550-551.

Ross, B. H., & Landauer, T. K. Memory for at least one of two items: Test and failure of several theories of spacing effects. *Journal of Verbal Learning and Verbal Behavior*, 1978, **17**, 669-680.

Smith, S. M. Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, **5**, 460-471.

Thorndike, E. L., & Lorge, I. *The teacher's wordbook of 30,000 words.* New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Underwood, B. J. Some correlates of item repetition in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 83-94.

Watkins, O. C., & Watkins, M. J. Buildup of proactive inhibition: A cue overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, **1**, 442-453.

NOTES

1. We used pairs for two reasons. First, having the subject concentrate on finding a relationship between the members of a pair should help to prevent the formation of associations between pairs. Theoretically, each pair is treated as a separate "unit," and the predictions hold for repeated units made up of pairs as well as for units made up of single words. Second,

we conducted a parallel experiment using cued recall. The cued recall data were problematic in that the overall level of recall was actually lower than that in free recall and the subjects' use of the cues to aid recall appeared to be influenced by spacing conditions. For these reasons, the cued recall data are not reported.

2. We thank Tom Landauer for suggesting this procedure.

3. A more complete example is as follows. The once-presented items, A and B, are presented in different global contexts (lists), and Item C is presented twice, once in each context. Assume that the encoded contextual elements are "abx" for Item A, "cdy" for Item B, and "abcdxy" in the unitary trace of Item C; "abcd" represent local components, and "xy" represent the global components. Assume that the probability of recall is a monotonic (but not necessarily linear) function (f) of the number of components shared by the trace and the cue.

After a short retention interval, assume that recall takes place in context "dey." Using that context as a cue, the probability of recall of A is f(0), the probability of recalling B is f(2), and for C, it is f(2). Now, suppose the subject attempts to recall the first list by mnemonically reinstating the global context present at that list, x. Using x as a cue, the probability of recall for A is f(1), for B, it is f(0), and for C, it is f(1). Therefore, recall of C using two independent retrieval attempts is $f(2) + f(1) - f(2) \times f(1)$. It is exactly equal to the probability of independently recalling A or B.

After a long retention interval, suppose recall takes place in context "fgz," but that the subject manages to reinstate the global contextual components x and y. The probability of recall for A is f(1), for B, it is f(1), and for C, it is f(2). Recall of C can be greater than predicted by independence for two reasons. First, f(2) may be larger than $f(1) + f(1) - f(1) \times f(1)$. Second, recall of A and B need not be independent. Given that the same components are used in recalling A and B, activation of A may inhibit activation of B.

(Received for publication February 21, 1980;
revision accepted May 1, 1980.)