Memory & Cognition
1973, Vol. 1, No. 4, 430-434

Spacing effects in picture memory*

DOUGLAS L. HINTZMAN and MIRIAM K. ROGERS
University of Oregon, Eugene, Oregon 97403

Three experiments were conducted to capitalize on the conclusion of Shaffer and Shiffrin (1972) that complex visual scenes are not rehearsed in testing the hypothesis that the effect of spacing on memory is due to rehearsal. In Experiment I, a list of vacation slides was presented in which both the number of repetitions and the spacing of repetitions were varied. Subsequent frequency judgments showed an effect of spacing much like that found using verbal materials. In Experiments II and III, effects of filled and unfilled spacing intervals were compared, and it was concluded that the spacing effect is primarily a function of the duration of the spacing interval. No evidence was found to support the notion that pictures are rehearsed. Rehearsal apparently cannot play the key role in an adequate, completely general explanation of the spacing effect.

The spacing effect—the fact that repetitions of an item massed closely in time lead to poorer performance on a later retention test than do repetitions that are spaced further apart—has been demonstrated using a variety of verbal materials, including nonsense syllables (e.g., Kintsch, 1966), words (e.g., Melton, 1967), and sentences (e.g., Underwood, 1970). In one explanation of the spacing effect, the process of rehearsal plays a critical role (Atkinson & Shiffrin, 1968; Greeno, 1967; Waugh, 1970). By rehearsal, we mean here the voluntary retrieval and reprocessing of a memory trace when the stimulus it represents is no longer physically present. According to the rehearsal hypothesis, an item for which the second presentation (P_2) occurs shortly after the first (P_1) is rehearsed less and, hence, is remembered less well on a later test than is an item for which the P_1-P_2 interval was long.

Direct support for the rehearsal hypothesis comes from the experiments of Rundus (1971), who studied the rehearsal patterns of Ss who were asked to rehearse aloud during presentation of a free recall list. When the spacing of repetitions was varied, less overt rehearsal was observed of words that had short P_1-P_2 intervals than of words that had long P_1-P_2 intervals. Rundus's data indicate, in addition, that the differential rehearsal took place entirely during the spacing interval. The number of rehearsals following P_2 did not depend on the P_1-P_2 lag. Thus, the spacing effect may have a rather trivial explanation: Longer spacing intervals may lead to better long-term retention simply because they give S more opportunities to rehearse P_1 before P_2 occurs.

Although the rehearsal hypothesis appears plausible, there is evidence against it. Bjork and Allen (1970) tested recall of word triplets. They interposed either a difficult or an easy task between P_1 and P_2 and always filled the P_2 test interval with a task of intermediate difficulty. Recall was not worse when the P_1-P_2 task was hard than when it was easy—in fact, it was somewhat better. Since the effect of differential rehearsal opportunities during the P_1-P_2 interval should favor the easy interpolated task over the difficult one, this somewhat paradoxical result poses a problem for the rehearsal explanation of the spacing effect.

The purpose of the first experiment reported here was to test the rehearsal hypothesis in another way, by using stimulus materials that Ss apparently do not rehearse. Shaffer and Shiffrin (1972) have reported that recognition memory for complex visual scenes is affected by stimulus "on" time but not by the blank "off" time following presentation; and from this, they have concluded that Ss do not (in fact, they cannot) rehearse such pictures. The stimulus materials chosen for the present studies, therefore, were color slides. The choice of dependent variable was dictated in part by the choice of stimuli. Description of complex pictures, like their rehearsal, is difficult, so recall would be almost impossible to score; and recognition memory for pictures is so good that any effects of spacing would likely be masked by a ceiling effect. Therefore, frequency judgments, which have no inherent ceiling and which show a strong spacing effect when verbal materials are used (Hintzman, 1969), were chosen as the dependent variable.

The conclusion that pictures are not rehearsed also led to their use in Experiments II and III, which had the purpose of determining whether it is the stimuli that are presented in the P_1-P_2 interval or the duration of the interval that causes the spacing effect. This problem is a difficult one to investigate using verbal materials, since any empty time can be used for rehearsal of either the immediately preceding stimulus or earlier stimuli. Not only are the temporal conditions of practice outside the E's precise control but also the independent manipulation of P_1-P_2 time vs P_1-P_2 items results in different degrees of learning, making the forms of spacing curves difficult to compare (e.g., Melton and Shulman data, reported in Melton, 1970). When pictures

*The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

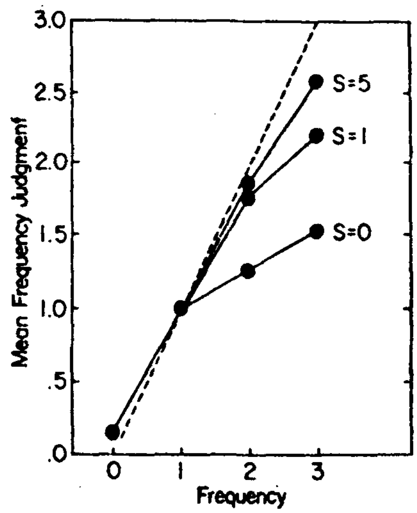

Fig. 1. Mean judged frequency as a function of frequency and the spacing of repetitions, Experiment I. (Pooled $S_{\bar{X}}$ = .060.)

are used, the presentation rate and degree of learning can be manipulated independently.

EXPERIMENT I

The purpose of this experiment was to determine whether the remembered frequency of occurrence of a picture is affected by the spacing of repetitions.

Method

Materials

The experimental stimuli were 120 color transparencies, taken from a collection of vacation slides and depicting scenes including landscapes, automobiles, people, buildings, etc. The pictures were selected to minimize interitem similarity. Thirty scenes were assigned at random to each of four frequencies of occurrence (F = 0, 1, 2, and 3) and an appropriate number of duplicates, including an additional test stimulus, were made of each. Within the F = 2 and F = 3 conditions, 10 scenes were assigned at random to each of three spacings (S = 0, 1, and 5 intervening items). For each F = 3 item, the P_1-P_2 and P_2-P_3 spacings were the same. In addition to the experimental slides, 10 filler slides were used.

The presentation sequence consisted of 190 slides, arranged in a continuous sequence in three Kodak carousel slide trays. The first five and last five slides were filler items. The 180 experimental slides were arranged, with appropriate spacings, in five overlapping blocks. All conditions except F = 0 (which occurred only on the test) were evenly distributed over the five blocks. Two additional sequences were constructed by rotating scenes in the F = 2 and F = 3 conditions across the three levels of spacing. Thus, there were three different rotations of the stimulus sequence presented to different Ss.

A single test sequence was used for all Ss. It included a test copy of each of the 120 experimental scenes. The sequence was arranged in five blocks, and each of the eight experimental conditions occurred equally often in each block.

Subjects and Procedure

The Ss were 27 paid volunteers obtained through the University of Oregon employment office. They were tested in three groups of up to 10 Ss each. A different rotation of items was presented to each group.

At the outset of the experiment, Ss were told that a series of pictures would be projected on the wall, that some of the pictures would be repeated, and that they were simply to study each picture for as long as it was presented and try to remember it for a later test. The nature of the memory test was not specified. The series of slides was then presented using a carousel projector paced by a timer at a 3-sec rate. The actual exposure duration was approximately 2.2 sec, and the blank interstimulus interval (ISI) was about .8 sec.

After the study series had been presented, Ss were shown the test slides. Each S was given a test form on which were 120 numbered blanks. They were told to write, in the blank appropriate to each test picture, the number of times they thought the picture had occurred in the preceding sequence. They were instructed to give a zero judgment for a picture if they thought it had not been presented before. The test slides were presented at a 5-sec rate.

Results

Mean judgments of frequency are presented in Fig. 1. The judgments increased as a function of both frequency and spacing, and the magnitude of the spacing effect was greater the more times the pictures occurred. For the purpose of comparison, a broken line representing perfect performance has been included in the figure. Frequency judgments of pictures appear, on the average, to be quite accurate. Accuracy was poor only when a picture was repeated at short spacings (S = 0 and 1).

In an analysis of variance using planned comparisons, linear trends of judged frequency on frequency and on spacing were both highly significant [$F(1,26) = 626.5$ and 90.5, respectively, both $ps < .001$]. The overall mean square error was 0.097.

The pattern of results is essentially the same as that obtained with words (Hintzman & Block, 1970; Underwood, 1969). The spacing effect obtained with pictures does not appear to be any more of less striking than that obtained using verbal materials. Thus, even if one is unwilling to admit that pictures cannot be rehearsed, as concluded by Shaffer and Shiffrin (1972) and will grant only that their rehearsal is inefficient compared to that of words, the magnitude of the obtained effect would appear to rule out an explanation in terms of rehearsal.

EXPERIMENT II

This experiment was designed to determine whether the spacing effect is a function of the duration of the spacing interval or of the number of items presented in the interval. Two sets of conditions, with spacings matched in terms of items but varied in terms of duration, were compared.

Method

The experiment was identical to Experiment I except for the following modifications:

(1) The presentation series, exclusive of the filler items at the beginning and end, was organized into 10 nonoverlapping blocks of 18 pictures each. Each frequency and spacing condition

except F = 0 occurred in each block. Only 10 of the 30 F = 1 items were considered experimental items. The other 20 were used as buffers, one at the beginning and one at the end of each block.

(2) In half the blocks, each picture was followed by a blank (opaque) slide. Thus, the ISI for these five blocks was 3.8 sec. In the other five blocks, there were no blank slides and the ISI was .8 sec, as in Experiment I. Except for the presence or absence of blank slides, the structures of the 3.8 ISI and .8 ISI blocks were identical. The two types of blocks occurred in alternating order: ABAB . . . to some Ss and BABA . . . to others. Several blanks were included among the fillers at the beginning and end of the entire series.

(3) Six different presentation sequences were used. To construct these lists, scenes in the F = 2 and F = 3 conditions were rotated across the three levels of spacing within their respective blocks, and each block of items occurred in both the 3.8-sec ISI and the .8-sec ISI conditions.

(4) There were 35 Ss, tested in six groups of up to 7 Ss each. A different stimulus sequence was presented to each group.

Results

Mean frequency judgments are presented in Table 1. As was the case in Experiment I, planned comparisons showed significant linear trends on both frequency and spacing [$F(1,34) = 743.9$ and 121.5, respectively, both ps $< .001$]. Judgments tended to be higher for repeated (F = 2 and F = 3) pictures from the 3.8 ISI blocks than for those from the .8 ISI blocks [$F(1,34) = 8.27$, $p < .001$]. Thus, P_1-P_2 time, with number of intervening items held constant, affected judged frequency. The interaction of Effects of Spacing by ISI was not significant [$F(1,34) = 2.91$, $p > .05$]. The overall mean square error for Experiment II was 0.142.

Was time the sole determinant of the effect of spacing? Two comparisons are directly relevant to this question. The .8 ISI pictures repeated at S = 1 and the 3.8 ISI pictures repeated at S = 0 both had a 3.8-sec spacing interval. In the former case, the interval contained another picture and, in the latter case, it contained a blank slide. Intervals containing pictures appear to have produced a slightly greater increment in judged frequency than empty intervals (1.62 vs 1.59 for F = 2 and 2.03 vs 1.83 for F = 3). The appropriate comparison, however, was not significant [$F(1,34) = 2.75$, $p > .05$].

Finally, the comparison of F = 1 judgments under the two ISI conditions confirms the recognition-memory

Table 1
Mean Frequency Judgments: Experiment II*

ISI	Spacing (Items)	Frequency 0	1	2	3
.8 Sec	0	0.11	1.06	1.40	1.67
	1	–	–	1.62	2.03
	5	–	–	1.87	2.50
3.8 Sec	0	–	1.02	1.59	1.83
	1	–	–	1.69	2.29
	5	–	–	1.82	2.66

*Pooled $S_{\bar{X}}$ = .064

Table 2
Mean Frequency Judgments: Experiment III*

Spacing Time (Sec)	Spacing Items	Frequency 0	1	2	3
.8	0	0.10	0.88	1.26	1.74
3.8	0	–	0.98	1.57	2.00
3.8	1	–	–	1.56	2.15
9.8	0	–	0.98	1.81	2.18
9.8	3	–	–	1.87	2.39

*Pooled $S_{\bar{X}}$ = .063

finding of Shaffer and Shiffrin (1972). Mean judged frequency of F = 1 pictures followed by an "off" time of .8 sec was about the same as that of pictures followed by a 3.8-sec "off" time. The difference was not significant, $F < 1$.

EXPERIMENT III

In Experiment II, only two direct comparisons could be made to determine whether other pictures presented during the spacing interval affected judged frequency. The purpose of this experiment was to compare four pairs of conditions in which spacings were matched in terms of time but the intervals were either filled or unfilled with other items.

Method

The method was the same as that of the previous two experiments, except for the following modifications:

(1) Spacing intervals were defined in terms of time rather than items. They were S = .8, 3.8, and 9.8 sec. The S = 3.8 and S = 9.8 intervals were filled either with pictures (one and three pictures, respectively) or with blank slides. In order that Ss would not be able to anticipate repetitions, P_2 of F = 2 and P_3 of F = 3 pictures were always followed by the same number of blank slides as had occurred in the preceding spacing intervals.

(2) There were three F = 1 conditions, in which the pictures were followed by zero, one, or three blank slides. Thus, the "off" times of .8, 3.8, and 9.8 sec matched the spacing intervals of repeated items.

(3) The presentation sequence, exclusive of the five filler items at the beginning and five at the end, was organized into five nonoverlapping blocks. Each condition was represented by one scene per block.

(4) Six sequences were constructed. Within their respective blocks, the F = 2 and F = 3 scenes were rotated among the 3 (spacings) by 2 (filled vs unfilled) conditions, and the F = 1 items were rotated through the three "off" times.

(5) There were 41 Ss, tested in six groups of up to 10 Ss each. A different rotation of the list was presented to each group.

Results

Mean frequency judgments are presented in Table 2. Again, planned linear trend tests showed that the increases produced by frequency and spacing were both highly significant [$F(1,40) = 731.3$ and 60.6, $p < .001$]. The effect of number of pictures in a spacing interval, tested as an interaction between the filled- vs

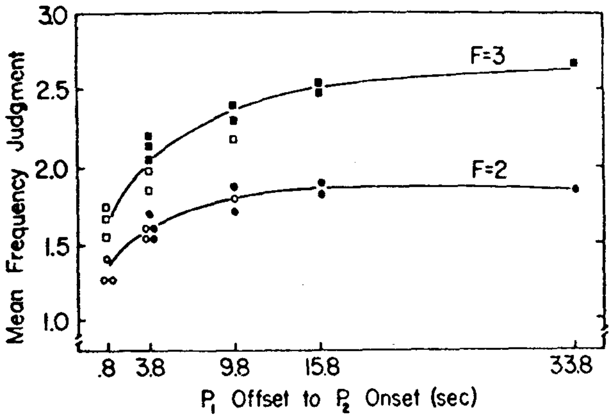

Fig. 2. Mean judged frequency as a function of P_1 offset to P_2 onset (and P_2 offset to P_3 onset) time. Data of all three experiments are combined. Circles: F = 2; squares: F = 3. Open data points represent unfilled spacing intervals.

unfilled-interval variable and spacing, was just barely significant [$F(1,40) = 5.24$, $p < .05$]. As in Experiment II, filled intervals led to slightly higher judgments than did empty intervals of the same duration. Also as in Experiment II, there was no effect of "off" time on judged frequency of F = 1 pictures. Neither the planned .8 vs 3.8 and 9.8 comparison nor the 3.8 vs 9.8 comparison was significant (both Fs near 1.0). The overall mean square error was 0.164.

DISCUSSION

The present results demonstrate clearly that a spacing effect, very much like that found using verbal materials, is obtained in memory for the frequency of occurrence of visual scenes. One could attempt to reconcile this outcome with the hypothesis that differential rehearsal is the cause of the spacing effect in two ways. One way would be to claim that the process underlying the effect of spacing found here is basically different from that at work in other situations. Logically, of course, it must be true that some of the mechanisms involved in the frequency judgment task are different from those involved in recall; likewise, the processes involved in pictorial memory and in verbal memory are not all identical. Therefore, the possibility that different processes may be producing superficially similar effects in different situations cannot be rejected in any final sense. Nevertheless, when one considers the striking similarity of the effects of spacing found with tasks as different as paired-associate recall and judgment of frequency and with materials as diverse as nonsense syllables, sentences, and pictures, it seems most parsimonious to assume, in the absence of clearly contradictory evidence, that the underlying cause is in all cases essentially the same.

If the assumption of a common underlying cause is accepted, then one might defend the rehearsal hypothesis by arguing that the conclusion of Shaffer and Shiffrin (1972), that complex visual scenes are not rehearsed, is false. In opposition to this second line of argument, two of the present results are in perfect agreement with the Shaffer and Shiffrin conclusion. First, Experiments II and III both confirm their finding that, for pictures presented once, duration of the "off" time following presentation has no effect on memory. If Ss did use this time to rehearse the previously presented picture, then the rehearsal had no observable lasting effect. Second, in both Experiment II and Experiment III, empty spacing intervals between repetitions led to slightly lower frequency judgments than did intervals of the same length that were filled with other pictures. If the cause of the spacing effect were greater rehearsal during long than during short spacing intervals, as the rehearsal hypothesis assumes, then the difference between filled and unfilled intervals would be expected to be large and in the opposite direction. Other than the spacing effect itself, then, the present data provide no evidence to suggest that the pictures were rehearsed.

If a spacing effect is found using stimulus materials that Ss do not rehearse, then the validity of the rehearsal explanation may be questioned for other situations as well. Earlier we mentioned the Bjork and Allen (1970) study, in which P_1-P_2 intervals were filled with rehearsal-preventing tasks of different difficulties, and the effect of the interpolated task on word recall was the opposite of that predicted by the rehearsal hypothesis. Taken together, the present experiments using pictures and those of Bjork and Allen using verbal materials suggest strongly that the rehearsal explanation of the spacing effect either is of very limited generality or is incorrect.

To the question of whether time or number of items determines the spacing effect, the answer offered by the present results is not quite so clear. In both Experiments II and III, the length of the spacing interval was a far more important determinant of judged frequency than was the number of pictures occurring during the interval. Nevertheless, the number of pictures did have an effect. Despite the small difference between filled and unfilled intervals and despite the fact that the total time for presentation was shorter in Experiment I than in Experiments II and III, the effect of spacing interval length was remarkably consistent across the three experiments. This can be seen in Fig. 2, in which data from all three experiments have been plotted as a function of time between repetitions. Empty spacing intervals are represented by open data points and filled intervals by solid data points. The smooth curves, fitted by eye, suggest that the spacing effect is almost entirely a function of time. This conclusion is quite acceptable for the F = 2 data (circles) but somewhat less so for F = 3 (squares), where a difference between filled and unfilled intervals is more apparent. One can speculate that if the rehearsal of verbal materials could be eliminated the effect of spacing on their retention might also prove to be primarily a function of time.

To summarize: the present experiments capitalized on the conclusion that complex pictures are not rehearsed, in order to test the hypothesis that the spacing effect is caused by rehearsal. The effect of the spacing of repetitions of pictures on their judged frequency was essentially the same as that found using verbal materials. Further, the effect of spacing on picture memory appears to be primarily a function of time and to be affected only slightly by events that occur during the spacing interval. Apparently, rehearsal–that is, the voluntary retrieval and reprocessing of a memory trace when the stimulus it represents is no longer present–cannot explain the present findings. If a completely general explanation of the spacing effect is to be found, it apparently will not be one in which the key role is played by rehearsal.

REFERENCES

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 2. New York: Academic Press, 1968. Pp. 89-195.

Bjork, R. A., & Allen, T. W. The spacing effect: Consolidation or differential encoding? Journal of Verbal Learning & Verbal Behavior, 1970, 9, 567-572.

Greeno, J. G. Paired-associate learning with short-term retention: Mathematical analysis and data regarding identification of parameters. Journal of Mathematical Psychology, 1967, 4, 430-472.

Hintzman, D. L. Apparent frequency as a function of frequency and the spacing of repetitions. Journal of Experimental Psychology, 1969, 80, 139-145.

Hintzman, D. L., & Block, R. A. Memory judgments and the effects of spacing. Journal of Verbal Learning & Verbal Behavior, 1970, 9, 561-566.

Kintsch, W. Recognition learning as a function of the length of the retention interval and changes in retention interval. Journal of Mathematical Psychology, 1966, 3, 412-433.

Melton, A. W. Repetition and retrieval from memory. Science, 1967, 158, 532.

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

Rundus, D. Analysis of rehearsal processes in free recall. Journal of Experimental Psychology, 1971, 89, 63-77.

Shaffer, W. O., & Shiffrin, R. M. Rehearsal and storage of visual information. Journal of Experimental Psychology, 1972, 92, 292-296.

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

Underwood, B. J. A breakdown of the total-time law in free-recall learning. Journal of Verbal Learning & Verbal Behavior, 1970, 9, 573-580.

Waugh, N. C. On the effective duration of a repeated word. Journal of Verbal Learning & Verbal Behavior, 1970, 9, 587-595.

(Received for publication March 17, 1973; revision received April 23, 1973.)