Memory & Cognition
1975, Vol. 3 (3), 287-294

What causes the spacing effect? Some effects of repetition, duration, and spacing on memory for pictures

DOUGLAS L. HINTZMAN, JEFFERY J. SUMMERS, and RICHARD A. BLOCK
University of Oregon, Eugene, Oregon 97403

Three experiments examined effects of the spacing of repetitions on memory for pictures. In Experiment I, the duration of the first presentation (P_1) was manipulated, as was P_1-P_2 spacing. The effect of spacing on judged frequency was independent of P_1 duration. In Experiment II, pictures were given M massed presentations just prior to the P_M-P_{M+1} spacing interval. The form of the spacing curve was independent of M. Neither experiment confirmed the prediction of "overhabituation," derived from the habituation-recovery explanation of the spacing effect. In Experiment III, subjects made both duration and frequency judgments. The duration judgment results were not consistent with the notion that subjects remember multiple massed presentations as single occurrences of especially long duration. Some evidence from Experiments I and III suggests that an interrupted stimulus is recognized better than one that is not interrupted—a finding that, if replicable, would support the habituation-recovery account of the spacing effect.

Why do two successive presentations of a to-be-remembered stimulus lead to poorer long-term retention than two presentations that are spaced further apart? A number of explanations of this puzzling finding have been proposed (cf. Hintzman, 1974). This paper reports experiments that were done to test two explanations of the spacing effect. One of these explanations has been proposed before, and the other apparently is new.

Briefly, the spacing effect can be characterized as follows: Let us call the first and second presentations of an item P_1 and P_2, respectively. As the P_1-P_2 interval is lengthened, long-term retention of the item increases, approaching an asymptote at a spacing interval of approximately 15 sec. The effect has been obtained in memory for words, nonsense syllables, sentences, and pictures; in tasks as diverse as paired associates, free recall, recognition memory, verbal discrimination, and the distractor task; using measures of recall and recognition accuracy, recognition latency, and judged frequency. When pictures are used as stimuli, empty and filled P_1-P_2 intervals are found to have about the same effect, suggesting that the process underlying the spacing effect is primarily a time-dependent one (Hintzman & Rogers, 1973). And the evidence suggests that it is deficient retention of P_2 rather than P_1 that produces the poor performance when the P_1-P_2 interval is short (Bjork & Allen, 1970; Hintzman, Block, & Summers, 1973).

The present experiments follow those of Hintzman and Rogers (1973) in using scenic pictures as stimuli and judged frequency as the primary dependent variable. Complex scenes have several advantages for experiments such as the present ones: They are easily remembered, they maintain subject's interest through long presentation sequences, and they are apparently difficult or impossible to rehearse (Cohen, 1973; Shaffer & Shiffrin, 1972). This last property makes them especially useful in investigating the spacing effect, because it gives the experimenter greater control over temporal variables than can be obtained when verbal materials are used.

HABITUATION-RECOVERY HYPOTHESIS

The first explanation of the spacing effect considered here is the habituation-recovery hypothesis (Hintzman, 1974). According to this hypothesis, whenever a stimulus item is studied some internal process involved in storing a memory trace of that stimulus becomes adapted—that is, its threshold of activation is raised. This adaptation or habituation continues for as long as attention is directed to the stimulus. When the stimulus ceases or attention is directed elsewhere, recovery from habituation begins. If P_2 occurs before recovery from the habituation to P_1 is complete, encoding of P_2 will be less effective than if P_2 is delayed. Thus the spacing effect occurs because immediately after a trace of P_1 has been encoded the pathway involved is in a refractory state. In order to be maximally effective, P_2 must occur outside the refractory period. Viewed in this way, the spacing curve is seen as tracing out the time-course of recovery from habituation.

If the spacing effect is due to habituation, then it may be possible to affect the time-course of recovery by producing different degrees of habituation prior to the

This research was supported by a research grant from the Office of Education, United States Department of Health, Education, and Welfare, and by Grant GB-40360 from the National Science Foundation. Requests for reprints should be sent to Douglas L. Hintzman, Psychology Department, University of Oregon, Eugene, Oregon 97403.

spacing interval (cf. "overhabituation" or "below-zero habituation" in the animal learning literature). In Experiment I, this was attempted by manipulating the exposure duration of P_1 and varying the P_1-P_2 spacing interval. In Experiment II, either one, two, or three massed presentations, followed by a varied spacing interval, preceded the final presentation of the picture. In both experiments it was predicted that the greater was the exposure of the stimulus just prior to the spacing interval, the slower would be the recovery traced out by the spacing curve. Thus a particular form of interaction was predicted in which the rate at which the spacing curve approaches asymptote is inversely related to the amount of exposure to the stimulus immediately preceding the spacing interval.

EXPERIMENT I

Method

Materials and design. The experimental items were 144 color pictures of a variety of vacation scenes. Each picture was randomly assigned to 1 of 22 experimental conditions. The 18 pictures in one condition did not occur during the presentation series, but served as distractor items on the frequency-judgment test. Another 18 pictures occurred one time during the presentation series, 6 at each of three different levels of exposure duration. The remaining 108 pictures occurred two times each. Six were assigned to each of 18 conditions representing the orthogonal combination of three levels of P_1 duration and six levels of P_1-P_2 spacing. For both single- and double-presentation pictures, the three levels of duration were set at 3, 6, or 9 sec, onset-to-onset time. Since the time required for slide changes was 0.8 sec, the actual exposure durations were 2.2, 5.2, and 8.2 sec. For double-presentation pictures, the six levels of spacing interval between presentations were 0.8, 3.8, 6.8, 9.8, 12.8, and 18.8 sec from offset of P_1 to onset of P_2. The P_2 duration of repeated pictures was always 5.2 sec (onset to offset).

The presentation series was divided into six blocks, and each of the 21 conditions (distractor condition not included) was represented by one picture in each block. The within-block order of conditions was random, subject to limitations imposed by spacing requirements. In each block, an additional two pictures, drawn from the same pool, occurred three times each. These pictures, which were not tested, were repeated at spacing intervals which varied. The first 10 and last 10 slides in the entire series were also filler items, drawn from the same pool, and some of these were repetitions. Altogether, there were 290 slides (including repetitions and fillers) in the presentation series.

The frequency-judgment test series, which was composed of one slide of each of the 144 experimental pictures, was divided into six blocks of 24 pictures. Each experimental condition was represented by one picture in each block, except for the distractor condition, which was represented by three in each block. The within-block order of conditions was random, and all blocks of the presentation series were represented about equally in each block. The test sheet for frequency judgments consisted of 144 consecutively numbered blank spaces.

Subjects and procedure. The subjects were 63 paid volunteers obtained through an advertisement in the University of Oregon campus newspaper. They were tested in six sessions of 9 to 12 subjects each. Across sessions, pictures were partially rotated through conditions, so that each picture occurred twice at each level of duration of the first (or only) presentation, and double-presentation pictures occurred once at each level of P_1-P_2 spacing.

At the start of each session, subjects were told that a series of 290 slides would be shown, that the series would take 29 min, and that the slides would appear for varying amounts of time. They were told that many of the pictures would be repeated, and that they were simply to study each slide for as long as it appeared and try to remember it for a later test, the nature of which was not specified. The slide series was then presented on a wall of the experimental room by a Carousel projector controlled by a tape recorder and Kodak sound synchronizer. Presentation was paced by pulses which had been prerecorded on the tape by computer.

Following presentation of the slides, test sheets were distributed, and subjects were told that a test series of pictures would be shown one at a time, and that they were to judge how many times each picture had occurred in the previous series. They were told to guess when uncertain, and if they did not remember a picture as having occurred, to give a judgment of zero. The test series was then presented at a 9-sec rate. The number of each test slide (1-144) was announced, so that subjects could write each judgment in the appropriate space.

Results

The mean frequency judgment for distractor (frequency of zero) pictures was 0.23. Mean judgments for the three single-presentation conditions and the 18 double-presentation conditions are presented in Figure 1.

Four features of the data should be noted. First, judged frequency of single-presentation pictures increased slightly with exposure duration. Second, the judgments given to double-presentation pictures tended to be higher the longer was the duration of P_1. This outcome is contrary to what would be expected if, as the habituation-recovery hypothesis suggests, the inhibitory effect of P_1 is greater the longer the subject has attended to P_1. Third, a spacing effect was obtained. And fourth, also contrary to the prediction of "overhabituation," the magnitude and form of the spacing effect were unaffected by the duration of P_1.

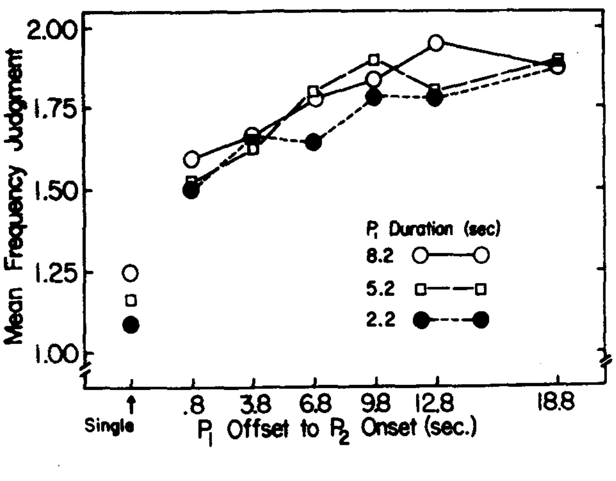

Figure 1. Effects of P_1 duration and P_1-P_2 spacing interval on judged frequency, Experiment I.

The preceding conclusions were all supported by a planned-comparisons analysis of variance (Grant, 1954). The analysis showed an effect of frequency, $F(1,62) = 1219$, $p < .001$; an effect of duration of single-presentation items, $F(1,62) = 7.50$, $p < .01$; and an effect of P_1 duration for double-presentation items, $F(1,62) = 10.1$, $p < .005$. To test for the spacing effect, comparison coefficients were selected which approximated the form of the spacing curve obtained in previous experiments using pictures (Hintzman & Rogers, 1973, Figure 2). For spacings of 0.8, 3.8, 6.8, 9.8, 12.8, and 18.8 sec, respectively, the coefficients were: −16, −6, 0, 5, 8, and 9. The overall effect of spacing was significant, $F(1,62) = 111$, $p < .001$; but three tests for pairwise interactions of P_1 duration with spacing were all nonsignificant, $F < 1$.

The outcome is clear. If the spacing effect is produced by habituation, then increases in habituation that will affect the spacing function apparently cannot be produced by extending the duration of P_1 beyond 2.2 sec. There remains the possibility, however, that repeated presentation of a picture might produce "overhabituation" even though increased duration does not. To test this prediction, we performed Experiment II.

EXPERIMENT II

Method

Materials and design. The experimental items were 140 color vacation pictures. Each picture was randomly assigned to 1 of 14 conditions. The 28 pictures in one condition served as distractor items. Another 28 pictures occurred one time each during the presentation series. Eight-four pictures occurred either two, three, or four times during the presentation series, seven in each of 12 conditions representing the orthogonal combination of three levels of massed repetition of the first M presentations and four levels of P_M-P_{M+1} spacing. The three levels of M (the same picture shown M times in succession) were one, two, or three; the four levels of spacing between P_M (the last massed presentation) and P_{M+1} were zero, one, three, and seven intervening items (0.8, 3.8, 9.8, and 21.8 sec, respectively).

The presentation series was divided into seven blocks, and each of the 13 conditions (distractor condition not included) was represented in each block. The within-block order of conditions was random, subject to limitations imposed by spacing requirements. Each of the 12 multiple-presentation conditions was represented in each block by one picture, and the single presentation condition was represented by three or four pictures. In addition, there were three or four filler items in each block, occurring two, three, or four times at spacing intervals that varied. The first two and last two slides in the entire series were also filler items. Altogether, there were 324 slides in the presentation series.

The frequency-judgment test, which was composed of 144 pictures, was divided into four blocks of 36. Each of the 12 multiple-presentation conditions was represented by two pictures in three of the blocks and by one picture in another. Both the single-presentation condition and distractor condition were represented by seven pictures per block, and the remaining test items were fillers. The within-block order of conditions was random, and all blocks of the presentation series were about equally represented in each block of the test series.

Subjects and procedure. The subjects were 41 paid volunteers obtained as in Experiment I. They were run in four sessions of 10 or 11 subjects each. Across sessions, pictures were rotated through spacings, but they remained in the same frequency condition and block. Thus each multiple-presentation picture occurred at each level of P_M-P_{M+1} spacing.

At the outset of each session, subjects were told that a series of 324 pictures would be shown for 3 sec each, that the series would take 16 min, that many of the pictures would be repeated several times, and that they were simply to study each slide and try to remember it for a later test. Presentation and test procedures were essentially the same as in Experiment I.

Results

The mean frequency judgment for distractor items was 0.17. Mean judgments from the other conditions are presented as a function of spacing and M, the number of massed repetitions preceding the spacing interval, in Figure 2. As is clear from the figure, mean judged frequency increased with M (as it should, since true frequency is M + 1), and also increased with spacing, at all levels of M. Further, the form of the spacing curve does not appear to have been altered systematically by changes in M. It is puzzling that the curve for M = 2 is closer to that for M = 1 than to that for M = 3 since frequency discrimination is typically better between lower than between higher frequencies (Hintzman, 1969). This outcome may reflect an inadvertent item selection problem: Pictures were rotated through all spacings within frequency levels, but not across frequencies. Nevertheless, the crucial rotation is the one that was carried out, since the prediction being tested concerns differences in the form of the spacing curves.

An analysis of variance confirmed the foregoing conclusions. Linear trend on frequency was significant, $F(1,40) = 1379$, $p < .001$; as was the overall effect of spacing, tested using comparison coefficients −6, −1, +3, and +4, $F(1,40) = 235$, $p < .001$. Three tests for interactions of M and spacing, involving all pairwise

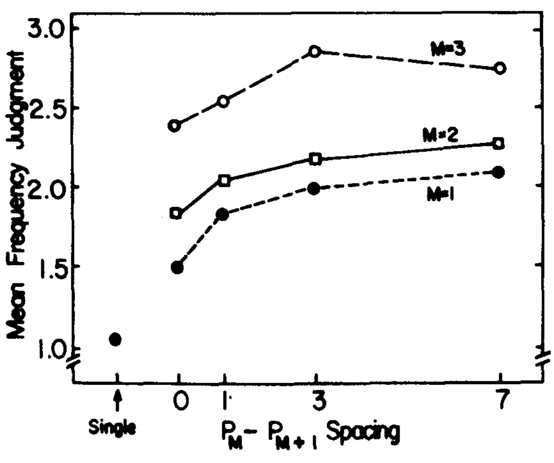

Figure 2. Effects of number of prior massed presentations (M) and P_M-P_{M+1} spacing interval on judged frequency, Experiment II.

combinations of the three levels of M, all failed to reach significance, $p > .05$.

These results are in complete agreement with those of Experiment I. If the spacing effect is caused by a habituation-recovery process, then apparently neither increasing the exposure duration of a stimulus beyond 2.2 sec nor repeating it several times in succession has any appreciable effect on the course of recovery from habituation. The prediction of "overhabituation" has not been confirmed.

INTERVAL RETRIEVAL HYPOTHESIS

Experiments I and II, in showing that neither increased exposure nor repetition just prior to the spacing interval influences the form of the spacing effect, reveal still another finding that may be important. The spacing effect appears to be a function of the time from offset of P_1 to onset of P_2, not of onset-to-onset time. This outcome is consistent with a hypothetical recovery process only if it is assumed that recovery is independent of the amount of prior exposure. It is also consistent with an explanation of the spacing effect that we shall call the interval retrieval hypothesis. This hypothesis makes a rather unorthodox assumption about the nature of the information that is encoded in memory, and so some space must be devoted to developing the argument.

The shortest P_1-P_2 interval used in the previous experiments was 0.8 sec. Suppose this interval were made successively shorter, until the offset of the stimulus was so brief that it was difficult to detect. Under such circumstances, it would not be surprising if subjects sometimes reported seeing the stimulus only once, since in order to know it occurred twice one would have to be aware of the interval separating P_1 and P_2. Now extend this notion to intervals of 0.8 sec and longer. Judging from memory that an item occurred twice instead of only once can be thought of as equivalent to remembering that an interval separated P_1 and P_2. Of course 0.8 sec of time is not likely to go unnoticed when it occurs. But it is possible that the memory trace of the P_1-P_2 interval is retrieved more easily the longer was the duration of the interval, just as the memory trace of a picture is retrieved more easily the longer was its exposure duration. This hypothesis requires reversal of the figure-ground relation customarily used in thinking about memory experiments. It is assumed that the retrieved information from which the subject infers that a stimulus occurred twice is the memory trace of the offset-to-onset interval rather than the trace of the stimulus item itself.

A difficulty with the interval retrieval hypothesis, in addition to its unorthodox nature, is that it does not lend itself as easily to explaining the effect of spacing on recognition or recall as it does to judged frequency. Nevertheless, because of the agreement of the hypothesis with the outcomes of Experiments I and II an attempt was made to test the hypothesis directly.

If subjects remember a massed-repetition item as having occurred only once because they fail to retrieve the spacing interval, then they should judge the exposure duration of the remembered "single presentation" as having been especially long. Experiment III tested this prediction using a task in which subjects judged both frequency and exposure duration. Since exposure duration must be varied in order for a request for duration judgments to seem reasonable, the presentation series included not only an orthogonal manipulation of frequency and spacing, but also an orthogonal manipulation of frequency and exposure duration. The latter manipulation also allowed a replication of a previous finding of relative independence of judgments of frequency and duration (Hintzman, 1970), this time using pictures instead of words.

EXPERIMENT III

Method

Materials and design. The experimental items were color transparencies of 96 vacation scenes. Each picture was randomly assigned to 1 of 14 conditions. The 18 pictures in one condition served as distractor items. Another 18 pictures occurred one time each, six at each of three levels of exposure duration (2.2, 5.2, and 8.2 sec). Thirty pictures occurred two times each, six in each of five conditions. In three of these conditions, exposure duration was 2.2 sec, and the P_1-P_2 spacing was either 0.8, 3.8, or 15.8 sec. In the other two, spacing was 15.8 sec and exposure duration was either 5.2 or 8.2 sec, with P_1 and P_2 durations the same for a given picture. A further 30 pictures occurred three times each, six in each of five conditions varying in spacing and exposure duration exactly as in the case of the double-presentation pictures. The nature of the incomplete orthogonal design is illustrated in Figure 3. For pictures that occurred two and three times, spacing was varied only for items having an exposure duration of 2.2 sec, and for items that varied in exposure duration, spacing was always 15.8 sec. The interval from the offset of one slide to onset of the next was always 0.8 sec.

The presentation series was divided into six blocks, and each of the 13 conditions shown in Figure 3 was represented by one picture in each block. The within-block order of conditions was random, subject to the limitations imposed by the spacing requirements. The first 14 and last 10 slides in the entire series were filler items, some of which were repetitions, and an additional 19 filler items were distributed throughout the presentation series. Altogether, there were 211 slides in the presentation series.

The test series, composed of the 96 experimental pictures, was divided into six blocks of 16 slides each. Each of the 13 experimental conditions was represented by one picture and the distractor condition by three pictures in each block. The within-block order of conditions was random, and all blocks of the presentation series were about equally represented in each block of the test series.

There were two similar test sheets for each subject, the first numbered 1-48 and the second 49-96. To the right of each number were (1) a column containing a 0, headed "Did not occur," (2) a column containing the numerals 1, 2, and 3, headed "Number of times," and (3) a column containing the numerals, 3, 6, and 9, headed "Exposure duration." A main heading of "Did occur" extended over both the second and third columns.

Subjects and procedure. There were 82 subjects, obtained as in Experiments I and II. They were run in 10 sessions of 7 to 10 subjects each. Across sessions, pictures that occurred two and three times were rotated through all 10 combinations of two or three presentations, the three levels of spacing, and the three exposure durations. The distractor and single-presentation pictures were rotated through these two conditions and the three levels of exposure duration.

The presentation instructions and procedure were essentially the same as in Experiment I. For the test, subjects were asked to judge how many times each test picture had previously occurred and for how long it had been shown, guessing if uncertain. If they did not remember a picture as having previously occurred, subjects were told to circle the 0 on the test sheet under the column headed "Did not occur." If they thought that a picture had occurred in the previous series they were to indicate how many times it was shown by circling either the 1, 2, or 3 in the column headed "Number of times," and the amount of time the picture was shown each time it occurred by circling either, 3, 6, or 9 in the column headed "Exposure duration." It was emphasized that every time a picture was shown it occurred for the same amount of time, and that the exposure duration judgment referred to each individual presentation, disregarding the number of repetitions and total exposure time.

Results and Discussion

Recognition proportions were first determined for each subject in each condition, by counting "Did not occur" responses as recognition failures. Then mean judgments of frequency and duration were determined for each subject, counting recognized items only. Thus judgments of zero frequency were not included in the frequency-judgment distributions as they were in Experiments I and II.

The proportion of false alarms made to distractor items was .09. For the other conditions, the mean proportions of pictures correctly recognized as old are presented in Figure 4. The left panel of the figure shows effects of frequency and spacing with duration constant at 2.2 sec, and the right panel shows the joint effects of frequency and duration with spacing constant at 15.8 sec. (Note that two data points in the left panel are duplicated in the right panel since the conditions were the same.)

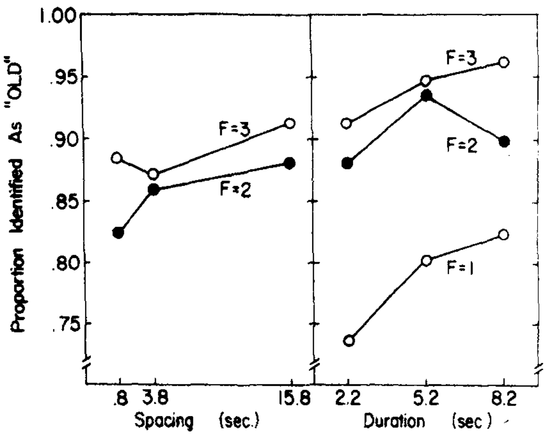

Figure 4. Recognition memory as a function of spacing, frequency, and exposure duration.

The data in the right panel clearly show that both frequency and exposure duration affected recognition performance. Linear trends on frequency and duration were both reliable, $F(1,81) = 141$ and 20.4, respectively, both $p < .001$. The data in the left panel show a spacing effect in recognition memory for pictures. Although the curve for frequency of three was somewhat irregular, the overall spacing effect, tested using comparison coefficients of −3, +1, +2 for spacings of 0.8, 3.8, and 15.8 sec, respectively, was significant, $F(1,81) = 5.72$, $p < .05$.

Mean judgments of frequency, conditionalized on recognition, are presented in Figure 5. The right panel shows that judged frequency increased with both frequency and exposure duration, although frequency had by far the greater effect. Linear trends on frequency and duration were both reliable, $F(1,81) = 992$ and 33.5, respectively, both $p < .001$. The effect of spacing, shown in the left panel, was also significant, $F(1,81) = 89.5$, $p < .001$.

Mean judgments of duration are presented in Figure 6. The right panel shows that judgments of duration, while an increasing function of both frequency and duration, were influenced more by duration than by frequency. Linear trends of duration judgments on both frequency and duration were significant, $F(1,81) = 40.8$ and 173.5, respectively, both $p < .001$. The finding that subjects were able to reliably judge duration indicates that the experiment is a valid test of the prediction regarding effects of spacing on remembered duration.

According to the interval retrieval hypothesis the subject mistakenly judges a massed-repetition item as having occurred once because of a failure to retrieve the trace of the P_1-P_2 interval. Thus, two short presentations should be remembered as one long presentation; and so the function relating judged duration to spacing should be the inverse of that for

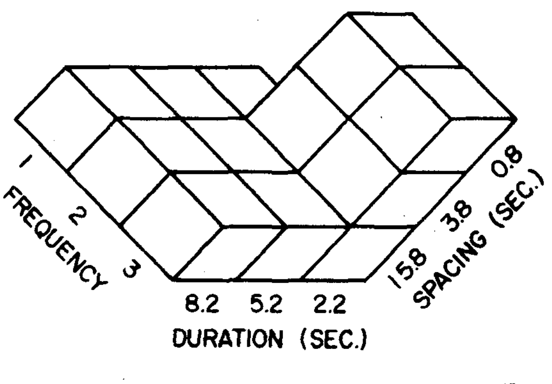

Figure 3. Design of Experiment III (6 pictures per cell).

judged frequency. As is shown in the left panel of Figure 6, this is not the case. The effect of spacing on judgments of duration, although slight, was significant and in the same direction as the effect on judged frequency, $F(1,81) = 12.5$, $p < .001$. Thus, overall mean duration judgments offer no support for the interval retrieval hypothesis.

Although it is reasonable to predict from the hypothesis that overall mean duration judgments will show a reverse spacing effect, a more sensitive test involves only those cases where judged frequency was below the true frequency. Three direct-difference t tests were done, selecting for each only the subjects who contributed scores to both categories being compared. In all three tests, the contrast was between the 0.8- and 3.8-sec spacing conditions, since it was between these two conditions that the greatest effect of spacing on judged frequency was found. For true frequency of two and judged frequency of one, 60 subjects contributed. Mean duration judgments for 0.8 and 3.8 sec spacings, respectively, were 5.04 and 4.80 $t(59) = 0.88$. For true frequency of three and judged frequency of one, 36 subjects contributed, and the respective means were 5.38 and 5.26 $t(35) = 0.37$. For true frequency of three and judged frequency of two, with 71 subjects contributing, mean duration judgments were 5.05 and 4.94, $t(70) = .48$. None of the three tests approached statistical significance Thus even when duration judgments were conditionalized on underestimation of frequency, there was no evidence that the 0.8-sec spacing interval resulted in longer remembered durations than the 3.8-sec interval. Discomfirmation of the predicted relationship between judged duration and spacing seems unequivocal.

The joint effects of frequency and duration on the three dependent variables, illustrated in the right-hand panels of Figures 4, 5, and 6, ae expressed in terms of explained variance in Table 1. This information can be compared with similar data from the Hintzman (1970) study in which words were used as stimuli. The comparison should be made with caution, however, as several details–most notably the range of exposure durations–differed in the two studies.

Note that while frequency and duration both contributed to recognition probability, total time accounted for no more than 69% of the variance, while the additive effects of the two variables accounted for as much as 93%. Thus a given amount of exposure time was less effective as one long presentation than it was as two or three spaced presentations. (When recognition was expressed as d′, the outcome was essentially the same.) More evidence relevant to this point will be given in the concluding section of the paper.

Frequency judgments, as Table 1 shows, were determined almost entirely by frequency. In contrast, duration judgments were affected much less by frequency than by duration. The additive effects of the two variabes on judged duration exceeded the effect of total time (97% vs. 91% of the variance). The overall pattern of frequency- and duration-judgment results is similar to that obtained using words (Hintzman, 1970).

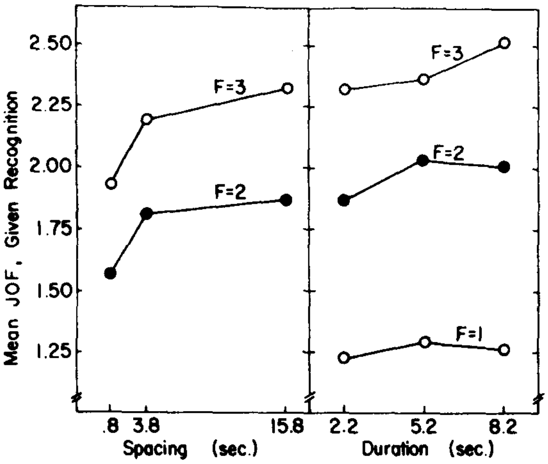

Figure 5. Mean judgment of frequency (conditional on recognition), as a function of spacing, frequency, and exposure duration.

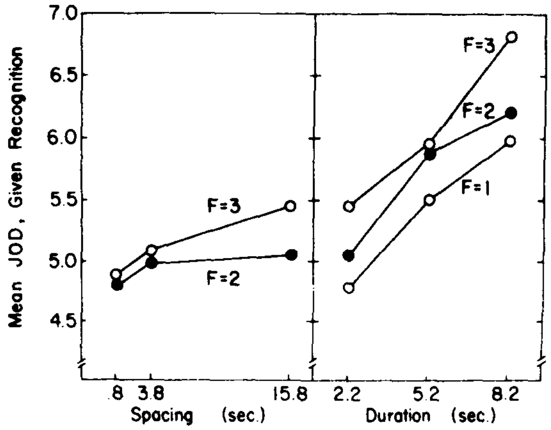

Figure 6. Mean judgment of duration (conditional on recognition), as a function of spacing, frequency, and exposure duration.

Table 1
Variance Among Means (r^2) Accounted for by Frequency, Duration, and Total Time

	Dependent Variable		
		Judgment Given Recognition	
Independent Variable	Recognition	JOF	JOD
Frequency (F)	.77	.96	.21
Log F	.83	.98	.20
Duration (D)	.09	.01	.76
Log F	.10	.01	.75
Total Time (F X D)	.54	.50	.86
Log F X D	.69	.51	.91

Note–JOF = judgment of frequency, JOD = judgment of duration.

However, the effect of duration on duration judgments was much stronger in the present experiment than in the previous one. This difference may be partly due to the different stimuli used in the two studies, although the wider range of exposure durations used here was undoubtedly a major factor. Taken together, the present frequency- and duration-judgment data confirm the conclusion of Hintzman (1970) that frequency and duration have effects on memory that are discriminably different to the subject at the time of retrieval.

CONCLUSIONS

Neither the prediction that "overhabituation" produced by prolonged or repeated exposure of a stimulus would affect the course of recovery, and hence the form of the spacing curve, nor the prediction that massed presentations would tend to be remembered as one long exposure was confirmed in the present experiments. Thus, both the habituation-recovery hypothesis and the interval retrieval hypothesis have failed to receive support in deliberate attempts to find evidence in their favor. We are seemingly left with only one new bit of information about the spacing effect: that it is a function of time from the offset of P_1 rather than from its onset.

The predictions of three alternative explanations of the spacing effect that were not considered in the previous discussion may be related to this finding. The first explanation, consolidation theory, is vague with respect to the relative importance of the onset and offset of P_1 in producing the effect. This hypothesis holds that the spacing effect is due to an interaction between the processing of P_2 and the ongoing processing or consolidation of P_1 in the period following its presentation (Landauer, 1969; Peterson, 1966). Since the theory does not address the question of effects of exposure duration, it is not clear whether consolidation should be assumed to start at the onset of P_1 or at its offset or whether exposure duration should affect the intensity and duration of the subsequent consolidation activity. Thus, different versions of consolidation theory would have predicted quite different outcomes of Experiment I.

A second explanation of the spacing effect, the contextual change hypothesis, begins with the assumption that elements of the stream of consciousness present when a stimulus occurs become associated with the representation of the stimulus, and provide information, at the time of retrieval, that the stimulus occurred in the experimental setting (Anderson & Bower, 1972). The spacing effect is explained by the additional assumption that massed repetitions tend to sample the same contextual elements, and so they aid retention less than do repetitions that are spaced further apart in time (Bower, 1972; Hintzman, 1974). The problem with this hypothesis, also, is its ambiguity with respect to effects of duration of P_1. Is the turnover of elements in the stream of consciousness strictly correlated with time, or are the same elements maintained somehow throughout the exposure of a given stimulus, from its onset to its offset? Again, different versions of the hypothesis would make different predictions regarding the duration manipulation of Experiment I.

A third hypothesis is that the encoding of P_2 is more effective the greater is the degree of forgetting of P_1. An extreme statement of this position, which holds that the subject will encode P_2 only if P_1 cannot be retrieved (Tzeng, 1973), is certainly incorrect, since the spacing effect occurs with materials for which recognition accuracy is very high. A weak version of the hypothesis, relating the effectiveness of encoding of P_2 to the relative difficulty of P_1 retrieval, is more plausible. However, neither the duration of P_1, in Experiment I, nor the number of prior presentations of the item, in Experiment II, influenced the form or magnitude of the spacing effect. Since both manipulations are known to affect recognition, their lack of an effect on the spacing curve appears inconsistent with a hypothesis that relates the spacing effect to retrieval difficulty.

Can anything positive be concluded about the nature of the process underlying the spacing effect? One unanticipated clue, which was discovered in the recognition data of Experiment III, is presented in Figure 7. At the left of each row in the figure is shown the temporal distribution of study time for pictures in a given condition. To the right of this are listed the total presentation time and the probability of recognition for that condition. Comparisons can be made between conditions that are matched with regard to initial onset to final offset times. The top panel compares one 5.2-sec exposure with two massed 2.2-sec exposures. The middle

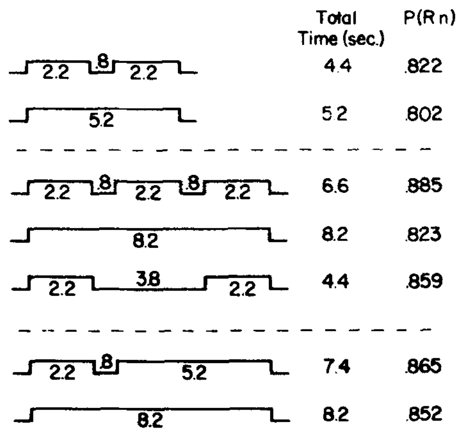

Figure 7. Effects of probability of recognition (right column) of selected conditions of Experiment III (top two panels) and Experiment I (bottom panel). Interrupted pictures are better recognized than comparable uninterrupted pictures.

panel compares one 8.2-sec exposure (in the center) with three massed 2.2-sec presentations and with two 2.2-sec presentations having an offset-to-onset time of 3.8 sec. In the bottom panel, a comparison of two appropriate conditions from Experiment I is made. In all comparisons, recognition memory for the interrupted stimulus was better than that for the continuous stimulus with which it is compared. This was true despite the fact that the total study time of the interrupted stimulus was less. While only the .885 vs. .823 comparison of the middle panel was statistically significant $t(81) = 3.11$, $p < .01$, the directions of all the differences favored the interrupted stimuli over the continuous ones.

Apparently, at least within certain limits, turning a picture off briefly one or more times produces better retention than simply leaving the picture on. This phenomenon might be called the interrupted stimulus effect. The present results do not allow us to eliminate possible trivial explanations of this effect (for example, onset of a picture may cause the subject to examine it briefly to see whether it has changed). Nevertheless this finding, together with the evidence that the joint effects of duration and frequency on recognition memory are not multiplicative (Table 1), seems entirely consistent with the spacing effect itself. It may be that the underlying process is the same. The facilitative effect of stimulus interruption suggests that, despite the lack of evidence of "overhabituation" in Experiments I and III, some sort of habituation-recovery hypothesis may yet turn out to be the best explanation of the effect of spacing on memory.

REFERENCES

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. Psychological Review, 1972, 79, 97-123.

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

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C: Winston, 1972. Pp. 85-123.

Cohen, G. How are pictures registered in memory? Quarterly Journal of Experimental Psychology, 1973, 25, 557-564.

Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. Psychological Bulletin, 1956, 53, 141-154.

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. Effects of repetition and exposure duration on memory. Journal of Experimental Psychology, 1970, 83, 435-444.

Hintzman, D. L., Block, R. A., & Summers, J. J. Modality tags and memory for repetitions: Locus of the spacing effect. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 229-239.

Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Potomac, Md: Lawrence Erlbaum Associates, 1974. Pg. 77-99.

Hintzman, D. L., & Rogers, M. K. Spacing effects in picture memory. Memory & Cognition, 1973, 1, 430-434.

Landauer, T. K. Reinforcement as consolidation. Psychological Review, 1969, 76, 82-96.

Peterson, L. R. Short-term verbal memory and learning. Psychological Review, 1966, 73, 193-207.

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

Tzeng, O. J. L. Stimulus meaningfulness, encoding variability, and the spacing effect. Journal of Experimental Psychology, 1973, 99, 162-166.

(Received for publication August 16, 1974; accepted September 4, 1974.)