Journal of Experimental Psychology: Human Learning and Memory
1975, Vol. 104, No. 1, 31–40

Spacing Judgments as an Index of Study-Phase Retrieval

Douglas L. Hintzman, Jeffery J. Summers, and Richard A. Block
University of Oregon

It is hypothesized that the ability of subjects to judge how far apart two presentations of a word were in a list reflects study-phase retrieval of the trace of the first presentation of the word by its second presentation. Experiment 1 supported this hypothesis by demonstrating that the accuracy of spacing judgments for associatively related pairs of words, like that for repeated words, was high compared to that for unrelated words. Experiment 2 used spacing judgments to measure retrieval upon repetition of a homograph. In three conditions, context words accompanying a homograph on its two presentations were either the same, biased the same meaning, or biased different meanings. In all three conditions, later spacing judgments were more accurate than in an unrelated-word control. Accuracy did not depend on whether the two context words biased the same meaning or different meanings of the homograph.

If a subject studies a word list in which some of the words occur twice, and later is shown one of the repeated words and asked to judge the spacing of presentations of the word in the original list, he is able to do so with some accuracy. Performance in a control condition, requiring a judgment of the spacing of two unrelated words that occurred one time each, is much poorer. This result was interpreted by Hintzman and Block (1973) in terms of study-phase retrieval. They assumed that the second presentation of a word (P_2) retrieves the trace of the first (P_1), and that this retrieval results in an implicit judgment of the recency of P_1. The implicit recency judgment itself is then stored in memory, and when it is retrieved on the later test, the subject can use it in making the required judgment of P_1–P_2 spacing.

If this interpretation is correct, then spacing judgments might be used in certain situations as an unobtrusive index of spontaneous study-phase retrieval. The distribution of spacing judgments in a given condition should consist of two components: judgments that are guesses, and those that are based on recency information encoded when retrieval took place. Thus, the degree to which the second of two events retrieves the trace of the first should be reflected in the accuracy of the mean judgments of their spacing, given on a later test. This paper reports two experiments that replicate the Hintzman and Block (1973) finding, confirm a prediction of this interpretation of spacing judgments, and demonstrate how spacing judgments can be employed as a tool to investigate spontaneous study-phase retrieval.

Experiment 1 compared judgments of spacing of three types of word pairs: *related* words (e.g., QUEEN–KING); *unrelated* words (e.g., SPIDER–TABLE); and repeated, or *same* words (e.g., WAR–WAR). It was assumed that the second member of an associatively related pair would tend to retrieve the trace of the first member, thus, according to the hypothesis, providing the necessary information for performance on the later spacing-judgment test. Two comparisons are important: (a) If spacing judgments given to pairs of related words are more accurate than those given to pairs of unrelated words, the study-phase retrieval hypothesis is supported. (b) If the hypothesis is accepted, then the degree to which related-word judgments approximate those given to same-

This study was supported by a research grant from the Office of Education, U.S. 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, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

word pairs is a measure of the reliability with which the second associate retrieves the trace of the first. This latter comparison assumes that a repetition of the word is the most effective retrieval cue of all.

Experiment 2 concerned the extent to which context words affect retrieval when the repeated word is a homograph. Four conditions were used, in which the relations between P_1 and P_2 of a homograph were manipulated by pairing with context words: *same pair* (e.g., POKER–DECK, POKER–DECK); *same meaning* (e.g., POKER–DECK, CARD–DECK); *different meaning* (e.g., POKER–DECK, SHIP–DECK); and *different pair* (e.g., POKER–DECK, CEREAL–CORN). The expectation was that judged spacing of homograph repetitions in the different-meaning condition would be less accurate than in the two meaning-preserving conditions, reflecting the importance of semantic context in retrieval.

Experiment 1

Method

Materials and design. A pool of 96 pairs of common English nouns was generated. The two members of each pair were judged by the experimenters to be strongly associated. An attempt was made to minimize associative relationships between members of different pairs. The types of associative relationships varied considerably. Examples of some are: WAR–PEACE, TABLE–CHAIR, KING–QUEEN, and SPIDER–WEB.

Using the original pool, a total of 72 experimental pairs of words were formed in the following way. Four pairs were randomly assigned to each of six same-word, six related-word, and six unrelated-words conditions. To form a same-word pair, a word from one of the original pairs was paired with itself on the spacing-judgment test (e.g., WAR–WAR). For each related-word pair, the original associative pairing was used on the test (e.g., CHAIR–TABLE). To form each unrelated-word pair, a word from one of the original pairs was randomly paired with a word from another of the original pairs (e.g., WEB–KING). In one of the six conditions under each of the three types of pairs, the two words occurred on the test sheet, but neither appeared in the word list. Thus the frequency of occurrence of the test pair in the list was zero ($F = 0$). In another condition, one member of the test pair appeared in the word list ($F = 1$). In four other conditions, both words appeared in the word list ($F = 2$), and the spacing between the two words was varied ($S = 0$, 3, 10, or 25 intervening items). When both words occurred and were related, the order of the two words was such that, intuitively, the second seemed more likely to elicit the first as a free associate (e.g., CHAIR preceded TABLE).

A slide of each experimental word was constructed by typing it on white paper and mounting it in a slide frame. Fifteen slides of filler words, selected from the same general population of words, were similarly constructed. The 123 slides were arranged in a continuous sequence in two slide trays; the first 10 and last 5 slides were filler words. The 108 slides of experimental words were divided into four partially overlapping blocks of 27 slides each, and all conditions except $F = 0$ were represented in each block. The order of conditions within each block was random, subject to the spacing requirements.

A single test form was used for all subjects. On it, the 72 word pairs were typed, with a blank line for the subject's spacing judgment appearing between the 2 words of each pair. In the related and unrelated conditions, the member of the pair that had appeared earlier in the list always appeared on the left. The order of the pairs on the test sheet was random with the restriction that each experimental condition was represented once and each block of the list was represented approximately four times in every block of 18 word pairs.

Subjects and procedure. The subjects were 55 paid volunteers obtained through an advertisement in the University of Oregon campus newspaper. They were tested in six groups of 6–12 subjects each. Between sessions, the particular word pairs of each type were rotated through the six frequency and spacing conditions.

The procedure was almost identical to that used by Hintzman and Block (1973). The subjects were told that a series of words would be presented at a 5-sec rate, that some of the words would be repeated, and that they were simply to study each word and try to remember it for a later test. The nature of the test was not specified. After presentation of the list, subjects were told to cross out all words on the test form that they did not remember, treating both members of same-word test pairs as different events. When they did remember both members they were instructed to estimate the number of other words that had occurred in between the two, restricting their judgments to the numbers 0 through 25. It was pointed out that when both members of a pair had been presented, the word on the left occurred earlier than the one on the right.

Results and Discussion

Recognition. On the test, subjects were to cross out individual members of the pairs that they did not recognize as having occurred in the list. As in the earlier study (Hintzman & Block, 1973), this procedure resulted in high rates both of hits and of false alarms. The proportions of pairs in each experimental condition for which subjects indicated that both members were old are given in Table 1. False alarm rates (F = 0 and F = 1 in the table) were about the same for the three types of test pairs. The same-word conditions show a typical spacing effect, with the probability of recognition increasing as spacing increased from 0 to 3 intervening items. The unrelated conditions appear to have been unaffected by spacing. The related conditions show a spacing effect that is the reverse of that found when the same word is repeated. That is, if associated words occurred together in the list, it was more likely that they would both be recognized than if they occurred far apart.

The effects of spacing found in Table 1 were tested for reliability using planned comparisons (Grant, 1956). Since the spacing effect typically asymptotes at a P_1–P_2 interval of about 15 sec (Hintzman, 1974), the comparison coefficients chosen for S = 0, 3, 10, and 25 were −3, +1, +1, and +1, respectively. Both the same-word spacing effect, F (1, 54) = 7.73, $p < .01$, and the related-word "reverse" spacing effect, F (1, 54) = 11.0, $p < .001$, were significant, while the unrelated word effect was not ($F < 1$). When the effects of spacing on the first two conditions were tested as interactions against the unrelated condition, the Same × Unrelated interaction was significant, F (1, 54) = 4.84, $p < .05$, while the Related × Unrelated interaction was not, F (1, 54) = 2.46, $p > .05$. It seems likely, however, that the latter test would have been significant if monotonic comparison coefficients had been used, since a similar effect of spacing of related items has been found in free recall by Glanzer (1969) and in recognition memory by Jacoby and Hendricks (1973). The Same × Related interaction was highly significant, F (1, 54) = 20.0, $p < .001$.

One characteristic of the present data does not replicate a finding of Jacoby and Hendricks (1973). They found that the first-occurring members of related pairs were recognized better than the members that occurred second. Hit rates for first and second members of our related pairs, collapsed across spacings, were .905 and .906, respectively. Both proportions decreased with spacing to about the same degree. Temporal order of occurrence was confounded in the present study with left–right order on the test sheet; and so this apparent discrepancy with the results of Jacoby and Hendricks could be due to response bias.

TABLE 1
PROPORTIONS OF TEST PAIRS BOTH JUDGED *Old* IN EXPERIMENT 1

Condition	Type of word pair		
	Same	Related	Unrelated
F = 0	.30	.31	.36
F = 1	.55	.46	.51
F = 2, S = 0	.78	.91	.85
S = 3	.87	.85	.83
S = 10	.86	.82	.83
S = 25	.86	.81	.82

Note. Abbreviations: F = frequency, S = spacing.

Spacing judgments. Judgments of spacing were given only when the subject indicated that both members of the pair had occurred. For all three types of word pairs, judged spacing was affected by the number of words in the pair that actually were old. Selecting data only from the 20 of 55 subjects who gave at least one spacing judgment in each of nine conditions (three word pair types × three frequencies, with F = 2 conditions collapsed over spacings), the following mean spacing judgments were obtained: for same-word pairs the F = 0, 1, and 2 values were 9.69, 6.61, and 5.62, respectively; the corresponding values for related-word pairs were 9.03, 8.69, and 7.34; and those for unrelated-word pairs were 9.79, 8.25, and 7.18. Linear trend on frequency was significant overall, F (1, 19) = 21.1, $p < .001$. There were no significant interactions among conditions. Since no spacing information should be available in memory in the F = 0 and F = 1 conditions, it seems likely that this decrease in spacing judgments across F = 0, 1, and 2 (which was also obtained by Hintzman and Block, 1973) reflects a bias to give high spacing judgments when guessing.

Of primary interest here are comparisons of spacing judgments among F = 2 conditi-

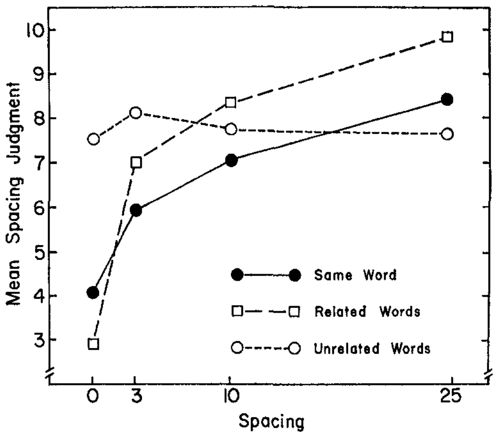

FIGURE 1. Mean judgment of spacing as a function of spacing for same, related, and unrelated conditions of Experiment 1.

tions only. For this analysis, so that within-subject analyses would be possible, data were used only from subjects who contributed at least one observation of each of the 12 proportions in the bottom four rows of Table 1. Of 55 subjects, there were 52 who met the criterion for inclusion, so little data was lost due to this selection. Means of subject means are presented in Figure 1.

The same-word and unrelated-word curves of Figure 1 essentially replicate the findings of Hintzman and Block (1973). Planned tests for linear trend confirm this conclusion. The increase in same-word judgments as a function of spacing was highly significant, F (1, 51) = 29.9, $p < .001$, while changes in unrelated-word judgments were not reliably affected by spacing ($F < 1$). The Same × Unrelated interaction was reliable, F (1, 51) = 25.3, $p < .001$. Mean judged spacing in the same-word condition exceeded that of the unrelated-word condition at the longest spacing, although the degree of crossover was not as great as in the earlier study.

The present prediction regarding the related-word curve was that it would increase monotonically with spacing, but be somewhat less steep than the same-word curve. As is apparent in Figure 1, the predicted monotonic increase was found, F (1, 51) = 64.7, $p < .001$. But the curve appears steeper than the same-word curve, and this is confirmed by the Same × Related interaction in the test for linear trend, F (1, 51) = 5.05, $p < .05$. This result requires further interpretation.

First, note that the monotonic increase in the related-word curve is exactly what was expected according to the hypothesis stated earlier—that reliable judgments of the spacing of two events reflect study-phase retrieval of the trace of the first event when the second event occurs. The strong effect of actual spacings on judged spacing of related words, as compared with unrelated words, thus confirms the hypothesis. The further expectation that the curve for related-word pairs would be less steep than that for same-word pairs was based on the reasonable assumption that the most reliable retrieval cue for the memory trace of a word would be a repetition of the word itself. This assumption could be incorrect. However, inspection of Figure 1 suggests that the interaction of the same and related conditions has two components: First, at S = 0, spacing judgments to related words were lower than those to same-word pairs. Second, at S = 3, 10, and 25, spacing judgments were higher to related words than to same-word pairs. The second difference very likely reflects a bias by subjects to treat same-word test pairs somewhat differently. As was noted earlier, judgments of spacing erroneously given to F = 1 pairs were lower for same-word pairs than for either related or unrelated pairs (6.61 vs. 8.69 and 8.25, respectively), and this difference is evidence for such a bias. If this argument is accepted, then the same-word curve in Figure 1 might be adjusted upward to roughly coincide at S = 10–25 with the related-word curve. This would leave the difference between the two curves at S = 0, magnified by the upward movement of the same-word curve, to be explained.

The correct explanation of the difference between same and related pairs at S = 0 probably lies in the quite different effects that spacing has on recognition memory in the two conditions. As Table 1 shows, a spacing of zero inhibits retention when the same word is repeated, but enhances it when the second word is associatively related to the first. It seems likely that this difference holds within the conditionalized data presented in Figure 1, as well. That is, on many of the S = 0 same pairs, subjects are not really very sure that the word occurred twice, and so the spacing-judgment data include a relatively high proportion of guesses, which (as the F = 0 and F = 1 data show) tend to fall toward the upper end of the judgment scale. In the case of S = 0 related pairs, since retention is enhanced, fewer of the judgments should be guesses. A rough check on the distributions of judgments in the S = 0 same and related conditions confirms this interpretation. The medians in both conditions were nearly identical (about .5), while the percentages of spacing judgments greater than 10 were different (14.9% in the same condition and 7.5% in the related condition). Thus it seems likely that the locus of the interaction between same and related judgments found in Figure 1 is at S = 0, and that different proportions of guesses, due to differential effects of massed presentations on retention, are the cause. As will be seen, a similar effect appeared in Experiment 2.

A final point to be made about the data in Figure 1 is that judged spacing increased from S = 10 (50 sec) to S = 25 (125 sec) in both the same and related conditions. Both increases were reliable, $F(1, 51) = 4.16$ and 5.13, respectively, both $ps < .05$. The increase for same-word pairs replicates that found by Hintzman and Block (1973). It seems unlikely, given this result, that contact of the second event with a hypothetical short-term memory trace of the first plays a crucial role in the encoding of information about spacing. More will be said on this point in the General Discussion section.

Experiment 2

Method

Materials and design. The experimental items were 80 three- and four-letter English homographs, each judged by the experimenters to have at least two completely different noun meanings (e.g., BANK). Four of the homographs were randomly assigned to each of four *same-pair,* four *same-meaning,* and four *different-meaning* conditions; the remaining 32 homographs were randomly paired, and four pairs were assigned to each of four *different-pair* conditions. The four conditions of each type varied in spacing (S = 0, 3, 10, or 25 intervening items) as in Experiment 1. Thus there were 16 experimental conditions (four pair types × four spacings), with four replications per condition for each subject.

Two slides represented each replication of each condition in the presentation series. On each slide, the to-be-remembered word was typed in capital letters and underlined. Above it, in lowercase letters, was a context word intended to bias a particular meaning of the homograph. The relationship between the two slides varied in the four types of conditions: in the same-pair condition, both the to-be-remembered word and the context word were the same on P_1 and P_2 (e.g., FLOWER–BULB, FLOWER–BULB). In the same-meaning condition, the to-be-remembered word was repeated, and different P_1 and P_2 context words biased the same meaning (e.g., FLOWER–BULB, TULIP–BULB). In the different-meaning condition, the homograph was repeated with context words intended to bias different meanings on P_1 and P_2 (e.g., FLOWER–BULB, LIGHT–BULB). In the different-pair condition, both the to-be-remembered word and the context word were different on the two slides (e.g., AIRPLANE–JET, WOOD–LOG).

The presentation series was divided into four partially overlapping blocks of 32 slides each, and all experimental conditions were represented in each block. The order of conditions in a block was random, subject to the predetermined spacing requirements. In addition, the first 5 and last 10 slides in the list were filler items, constructed in the same way as the experimental slides. Some of the fillers were repetitions. Altogether, the presentation series was 148 slides long.

A single test form was used for all subjects. On it were typed 88 word pairs. They included 48 same word pairs (e.g., BULB–BULB), using the homographs from the same-pair, same-meaning, and different-meaning conditions, and 16 different-word pairs using the homographs from the different-pair conditions. Context words did not appear on the test sheet. In addition to the F = 2 pairs, there were 12 F = 1 pairs (one member of the pair occurred in the list) of which 9 were same-word and 3 were different-word pairs, and 12 F = 0 pairs, of which 10 were same-word and 2 were different-word pairs. Words in these pairs were not rotated among conditions, and the data were not analyzed. The order of pairs on the test sheet was random with the restriction that each experimental condition was represented once and each block of the list was represented approximately four times in each block of 22 test pairs.

Subjects and procedure. There were 98 subjects, recruited in the same way as in Experiment 1. They were tested in 12 groups of 6–10 subjects each. Between sessions, the particular homographs assigned to the same-pair, same-meaning, and different-meaning conditions were rotated through these three conditions and the four levels of spac-

TABLE 2
PROPORTIONS OF TEST PAIRS BOTH JUDGED *Old* IN EXPERIMENT 2

Spacing	Condition			
	Same pair	Same meaning	Different meaning	Different pair
0	.60	.80	.85	.77
3	.77	.81	.82	.74
10	.77	.79	.81	.72
25	.80	.79	.78	.80

ing. The word pairs assigned to the different-pair conditions were rotated through the four spacings within that condition.

Instructions were similar to those of Experiment 1, with the nature of the final test unexplained. A pair consisting of a context word and a homograph was illustrated in an example slide, and the instructions, in part, were as follows:

> The function of the cue word is to make the meaning of the to-be-remembered word completely clear. Only the words in capital letters and underlined will be tested later. . . . We want you to study the words as follows: At the onset of a slide, read both of the words to yourself. Then use the remaining time to think about the underlined word and try to remember it for a later test. We are interested in how the use of cues affects memory. Therefore, please try to use the cue word as an aid to remembering the correct meaning of the to-be-remembered word. As the words are presented, you may notice that many of the words are repeated, as we are also interested in the effect of repetition on memory.

The presentation and test procedures were identical to those of Experiment 1.

Results and Discussion

Recognition. The proportions of times subjects indicated that both members of test pairs were old, by not crossing out either, are presented in Table 2. The effect of spacing on recognition differed, depending on the context manipulation. Spacing effects were tested for reliability using the same planned comparison coefficients that were used in Experiment 1. The same-pair condition replicated the spacing effect found with repeated words in Experiment 1, F (1, 97) = 40.3, $p < .001$. The different-meaning condition produced a marginally significant "reverse" spacing effect, F (1, 97) = 4.66, $p < .05$. Neither the same-meaning nor the different-pair conditions showed an effect of spacing (both Fs < 1).

Several investigators (e.g., Johnston, Coots, & Flickinger, 1972; Madigan, 1969) have found that changing the meaning of a homograph from P_1 to P_2 diminishes or completely eliminates the spacing effect in free recall. The present findings suggest that a reverse spacing effect, much like that obtained using associatively related words, can result from such a manipulation. The effect found here was only marginally significant, and the task used is quite different from free recall, so extrapolation to the free-recall situation can be done only with caution. Nevertheless, it seems possible that the lack of a spacing effect in free recall of homographs when their meanings are changed upon repetition does not necessarily reflect the independence of P_1 and P_2 in encoding and retrieval, as has been assumed. A lack of independence could have both positive and negative components that are additive and that vary with spacing.

The lack of a spacing effect, either positive or negative, in the same-meaning condition is puzzling. It may be that many of the P_2 context words biased the same meaning of a homograph, yet led to different encodings. For example, VIOLIN–CASE and GUITAR–CASE, JAZZ–BAND and MARCHING–BAND, and BEACH–BALL and TENNIS–BALL all represent cases in which the same "meaning" of the word is elicited on P_1 and P_2, but encodings in the form of visual images would differ. Thus the same-meaning spacing curve on Table 2, which is nearly flat across spacings, may actually be a combination of the same-pair and different-meaning curves.

Spacing judgments. Of the 98 subjects, 80 contributed at least one spacing judgment to each of the 12 experimental conditions. Means of mean spacing judgments for these 80 subjects are presented in Figure 2. The linear increase of judged spacing as a function of actual spacing was significant for the same-pair, same-meaning, and different-meaning conditions, F (1, 79) = 28.8, 49.2, and 45.5, respectively, all ps <

.001. For the different-pair condition, the effect was not significant ($F < 1$). Although none of the interactions among the three increasing curves were significant, it can be seen that in the same-pair condition, mean judgments did not change as much from $S = 0$ to $S = 3$ as they did in the same-meaning and different-meaning conditions. The effect is much like the Same × Related interaction found in Experiment 1. Increases in spacing judgments from $S = 10$ to $S = 25$ in the same-pair, same-meaning, and different-meaning conditions were all significant, F (1, 79) = 4.14, 4.10, and 5.93, respectively, all $ps < .05$. The interesting outcome concerns the same-meaning and different-meaning curves, which are nearly identical at all P_1–P_2 lags. On this matter, more will be said shortly.

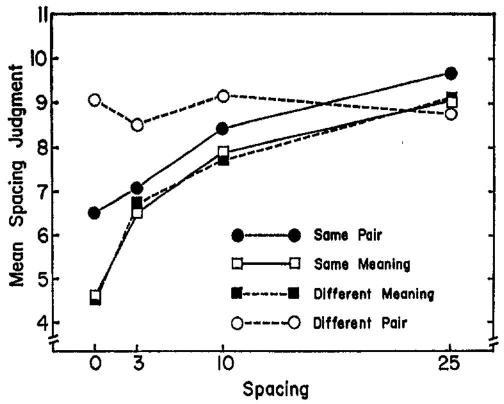

FIGURE 2. Mean judgment of spacing as a function of spacing for same-pair, same-meaning, different-meaning, and different-pair conditions of Experiment 2.

General Discussion

The present experiments were based on the hypothesis that the relative accuracy of subjects in judging the spacing of repetitions of a word, as compared to their inaccuracy in judging the spacing of single presentations of unrelated words, reflects the effect of study-phase retrieval of the trace of P_1 of the word when P_2 occurs. Experiment 1 confirms a strong prediction of this hypothesis by showing the subjects can judge the spacing of associatively related words. The following discussion assumes that the hypothesis is correct, and that one can therefore use judged spacing as an unobtrusive index of the spontaneous retrieval of the trace of one event by the occurrence of another during study of a list.

One problem to which the method is relevant concerns organization in free recall. Does the clustering of related words develop during study of the list, is it due to self-cuing at the time of retrieval, or does it involve some combination of the two? Evidence that organization develops during the study phase has been presented by Rundus (1971), who investigated rehearsal patterns for categorized lists. He found that when a word belonging to a particular category was presented, the words that tended to be rehearsed along with it were previously presented words of the same category. The present Experiment 1 suggests that the appearance of a related word in the rehearsal set may be a result of spontaneous retrieval of recently presented associates of the current word. The selection of the items to be rehearsed together may not be as much under the subject's voluntary control as has been assumed. Indeed, clustering in free recall occurs even when the subject does not anticipate a free-recall test; all that is necessary is that he be induced to process the words semantically (e.g., Hyde & Jenkins, 1969). It should be informative, therefore, to have subjects judge spacing of related words following performance in an incidental task requiring semantic processing, to determine whether study-phase retrieval of associates occurs even in the absence of the intent to learn.

Glanzer (1969) suggested that facilitation in the recall of related words does not take place unless the first word is still in the short-term store when the second occurs. Our recognition data, presented in Table 1, are consistent with Glanzer's free-recall data, and can be interpreted in the same way. Both the recognition and free-recall data, however, could equally well be explained by the hypothesis that presentation of a word primes or activates traces of related words in memory, so that their overt presentation a short time later leads to a stronger trace than would otherwise be established. Warren (1974) found evidence for such an activation process using a modified Stroop task. If the subject must hold a word in memory (e.g., BOY) while naming the color of ink in which an associate is printed (e.g., GIRL), color naming is slower than in a control condition in which the words are unrelated. Warren showed, in addition, that the degree of interference with color naming is a function of the strength of the forward (BOY–GIRL) rather than the backward association, suggesting that a priming process is indeed responsible.

As was mentioned earlier, the increase in mean judged spacing of related words from $S = 10$ to $S = 25$ suggests that the first word need not be in a short-term store or activated state for spontaneous retrieval to take place. This conclusion is supported by the work of Gruneberg (1972), who noted that subjects studying a word list detect associative relationships even when as many as 17 other items appear between the two related words. The experience was described by one subject as "like a bell ringing" (p. 279). The conclusion that retrieval of a related word does not depend on its presence in a hypothetical short-term store would explain why, in the rehearsal experiment of Rundus (1971), "presentation of a word from one of the categories not only increased the probability of other words from the category remaining in rehearsal, but perhaps more importantly, triggered the return to active rehearsal of words from that category which had been dropped from rehearsal" (p. 76). Items entered into the rehearsal set are not necessarily items already in the short-term store.

As Gruneberg (1972) noted, the question of whether presence in the short-term store is necessary for a facilitative effect on retention is not necessarily the same as that of whether it is necessary for detection of semantic relationships. The mechanism underlying the related word recognition data of Table 1 may be different from that underlying the related-word spacing-judgment data of Figure 1. Thus the possible role of a short-term store or activated state in the facilitative effect of presenting related words together in a list (Glanzer, 1968; Jacoby & Hendricks, 1973) is not disproved by evidence for related-word retrieval at intervals as long as 1 or 2 min, as obtained here.

Turning to Experiment 2, an argument can be made that is parallel to that just made concerning the role of a short-term store in retrieval of related words. The continued increase in mean judged spacing from $S = 10$ to $S = 25$ in the different-meaning condition of Experiment 2 seems to rule out the notion that P_1 must be in the short-term store at the time of P_2 in order for repetition of a homograph to be noticed. Based on the principle of encoding specificity (e.g., Tulving & Thomson, 1973), one would expect that the spacing-judgment curve for the different-meaning condition either would be less steep overall than the same-meaning curve, or else would asymptotically approach the different-pair curve sooner. Instead, the same-meaning and different-meaning curves were virtually identical. There are several possible explanations for this result. First, subjects may have paid no attention to the context words when assigning meanings to the to-be-remembered homographs. However, the effects of spacing on recognition, presented in Table 2, indicate that context words did affect performance differentially. And so it seems unlikely that subjects could have been simply ignoring the context words. Second, subjects may have somehow anticipated or discovered that many of the words had double meanings, and deliberately adopted a strategy of thinking of alternative meanings of each homograph in addition to the one biased by the context word. But this hypothesis does not really predict the obtained identity of the same-meaning and different-meaning curves, since the particular meaning biased by the context word on P_2 should be easier to retrieve than a meaning that is different, regardless of strategy. Perhaps the different-meaning curve would have been less steep, as the hypothesis suggests, if a presentation rate that allowed less time for such deliberate strategies had been used. A third possible interpretation is that the prior activation or priming of the P_1 meaning of the homograph increases the likelihood that it will be elicited by P_2, regardless of the context word that accompanies P_2. However, under the assumption that the activation declines within 30 sec or so, this hypothesis has similar difficulties with the fact that the different-meaning and same-meaning curves were identical at all spacings.

A final explanation is that presentation of a homograph automatically retrieves all of its meanings, in parallel, with context serving to select a particular meaning only after retrieval has taken place. In this case, since retrieval itself is independent of context, the match of one of the meanings retrieved by P_2 with a meaning stored for P_1 could be noticed by the subject just as readily when the meanings biased by the context words are the same as when they are different. In support of this interpretation, Conrad (1974) used the modified Stroop task to show that when a subject hears a sentence containing a homophone, meanings of the homophone in addition to the one biased by the semantic context of the sentence are contacted in memory. Following "The man drank water from the tumbler," for example, color-naming responses were slowed for both the test words *glass* and *gymnast*. It should be noted in addition that the retrieval of multiple meanings was reported many years ago by the introspectionists. Ogden (1917) had observers listen to a word, press a key when they were certain of the meaning, and then retrospect on the contents of consciousness during the response interval. Although his experiment was not directly concerned with ambiguities, many of his stimuli (e.g., FELT, PLAIN, WELL, BARK, EXECUTION) almost instantaneously—according to the reports—retrieved more meanings than one. So while the interpretation of the different-meaning result in terms of simultaneous retrieval of multiple meanings should not be uncritically accepted—particularly in the absence of information about how spacing judgments are affected by incidental learning instructions and manipulations of presentation rate—neither should it be rejected out of hand, as there is other support for such a view.

In their demonstration of encoding specificity, Tulving and Thomson (1973) showed that a word originally encoded with a weakly associated cue (e.g., CAVE–WET) is often not recognized as old when tested either in the presence of a strong cue (e.g., DRY–WET) or by itself—even though the subject is able to recall the word when the weak cue is present. This suggests the necessity of an intimate relationship between encoding context and retrieval context if retrieval is to succeed. Perhaps, given the apparent conflict between results of the present Experiment 2 and the principle of encoding specificity, we should consider abandoning our interpretation of spacing judgments in terms of study-phase retrieval. However, we have been unable to think of another interpretation that is comparable either in its intuitive appeal or its psychological plausibility. And it should be noted that Tulving and Thomson's subjects were pretrained on two successive practice lists, before the critical list was given, to expect a paired-associate recall test. This manipulation was not used here, and it may be necessary if powerful encoding specificity effects are to be produced. Thus the conflict between the encoding specificity principle and the results of Experiment 2 may be more apparent than real.

REFERENCES

Conrad, C. Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition,* 1974, *2,* 130–138.

Glanzer, M. Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior,* 1969, *8,* 105–111.

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

Gruneberg, M. M. The effect of distance between related words on the identification of semantic relationships. *Acta Psychologica,* 1972, *36,* 275–279.

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

Hintzman, D. L., & Block, R. A. Memory for the spacing of repetitions. *Journal of Experimental Psychology*, 1973, *99*, 70–74.

Hyde, T. S., & Jenkins, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 1969, *82*, 472–481.

Jacoby, L. L., & Hendricks, R. L. Recognition effects of study organization and test context. *Journal of Experimental Psychology*, 1973, *100*, 73–82.

Johnston, W. A., Coots, J. H., & Flickinger, R. G. Controlled semantic encoding and the effect of repetition lag on free recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 784–788.

Madigan, S. A. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 828–835.

Ogden, R. M. Some experiments on the consciousness of meaning. In, *Studies in psychology: Titchener commemorative volume*. Worchester, Mass.: Louis N. Wilson, 1917.

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

Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, *80*, 352–373.

Warren, R. E. Association, directionality, and stimulus encoding. *Journal of Experimental Psychology*, 1974, *102*, 151–158.

(Received April 1, 1974; revision received June 19, 1974)