A Reconsideration of the Extinction-Recovery Theory[1]

GEOFFREY KEPPEL[2]

University of California, Berkeley, California 94720

The extinction-recovery theory, which has been offered to account for the superior retention of paired-associate lists learned under distributed practice (DP), was tested in three experiments. These studies showed that (a) prior-list availability is unaffected by DP, (b) the DP effect is largely due to increased availability, and (c) a DP effect can occur under conditions of minimal interlist interference. These findings were considered severely damaging to the theory. The results of these and other studies were interpreted in terms of the operation of a selection process by which DP Ss are able to acquire a collection of relatively stable and strong associations.

Recent research has continued to support the conclusion that retention is largely a function of degree of learning and that variables which produce large differences in learning, e.g., meaningfulness, similarity, and individual differences, have little or no influence upon rate of forgetting (Underwood, 1964). One apparent exception to this conclusion is the facilitative effect on retention of distributed practice (DP) administered during learning. Specifically, retention of the last in a series of A-B, A-C, A-D, A-E lists (same stimuli but different responses in successive lists) has been found to be higher when the last list is learned under DP than under massed practice (Underwood, Keppel, and Schulz, 1962; Underwood and Schulz, 1961). Moreover, when the DP intervals are short (e.g., 1–3 min), this effect does not occur when the four lists form the A-Br paradigm, where the lists consist of four successive pairings of stimulus and response terms (Underwood et al., 1962). Finally, when the DP intervals are extremely long (e.g., 24 hr) the magnitude of the DP retention effect is impressive. For example, retention of the fourth list after one day was 89% for the DP condition and approximately 31% for two combined massed practice (MP) conditions (Keppel, 1964). While retention was almost zero (7%) for the two MP conditions after eight days, 72% and 34% retention was obtained for the DP condition after 8 and 29 days, respectively. These findings dramatically demonstrate that the massive interference effects of previously learned materials can be essentially eliminated for periods up to one month, when the critical list is learned with widely spaced DP intervals.

These results have been interpreted in terms of an extinction-recovery theory. Briefly, the theory consists of the following assumptions: (a) There must be high interlist interference, i.e., conflicting associations which will interfere with critical-list recall. (b) During fourth-list learning

there is extinction (or unlearning) of prior-list response terms. (c) These extinguished responses recover during the DP intervals and are re-extinguished on successive learning trials. (d) This process of successive extinction-recovery cycles produces a more permanent unlearning of the previously learned associations. (e) As a consequence, there is reduced interlist interference (and higher recall) for DP following a common retention interval for MP and DP. In addition to an explanation of the DP retention effect, the extinction theory has been able to account for other aspects of the data, e.g., inferior fourth-list learning, together with a higher incidence of interlist intrusions for DP (both findings presumably the result of interference from recovered associations), and more interlist intrusions for MP at recall (a reflection of differential interlist interference).

The main purpose of the present analysis was to provide tests of certain critical assumptions underlying the extinction-recovery theory. Such an analysis was considered necessary since there were already several potentially "soft" points in the theory to be found in the literature, where certain assumptions of the theory have not been supported by the data. (a) The conclusion offered by Underwood *et al.* (1962), that the DP effect will occur only when the interfering response terms are *not* present in the fourth list, is not supported by the data of Keppel (1964). In this study, relatively long DP intervals (24 hr) resulted in superior retention with the A-Br paradigm, where the specific response terms appear in all four lists. This discrepant finding may be handled by the theory through the assumption of differential recovery rates for various associations (Keppel, 1964). (b) The administration of DP on the third of a series of three A-B, A-C, A-D lists did not produce

greater unavailability, following third-list learning, of the two sets of prior-list responses (Keppel and Schwartz, 1965). This point may be considered noncritical, however, if it is assumed that three lists are not sufficient for producing the DP effect or that DP does not affect the level of prior-list extinction, but the rate of recovery. (c) At recall, Keppel (1964) also failed to find less prior-list recall for the DP condition. But even this finding is not severely damaging to the theory since prior-list recall was taken following the relearning of the fourth list, a procedure which may have eliminated any differences in prior-list availability present at the time of recall.

In the first experiment, the assumption of differential prior-list availability was tested directly by the administration of an unpaced recall test of the prior-list response terms at the time of recall. The expectation of the extinction theory, of course, is greater prior-list recall for the MP group. A second recall test, given to independent groups, involved the discrimination of the list membership for the four responses associated with the same stimulus in the four lists. This condition was included to test the hypothesis that the DP effect is the result of increased list differentiation for the DP group as the result of the differential treatment given to List 4 (see Houston, 1966; Houston and Reynolds, 1964). Specifically, this notion assumes that the change in experimental procedure (MP on Lists 1–3 and DP on List 4) makes List 4 sufficiently distinctive to increase list differentiation and thereby to reduce interlist interference for the DP condition. This hypothesis will be evaluated at a later time. For the present, however, it is important to note that both hypotheses predict better differentiation of List 4 for the DP group, i.e., the extinction-recovery notion might also specify less

critical-list confusion for DP as the result of reduced prior-list availability. Thus, while the differentiation test does not provide differential predictions, it does represent a critical test of the two hypotheses in the sense that a failure to find DP superiority would be damaging both to the extinction-recovery and differentiation explanations.

EXPERIMENT I

Method

Design. The design was a 2×3 factorial, the first factor consisting of a variation of the conditions of List-4 learning (MP or DP). The second factor consisted of three types of recall tests given 48 hr following training. The first two, a test of prior-list availability and a test of list differentiation, have already been mentioned. The third test consisted of the paced recall of List 4, a condition which was included to provide a replication of the original experiments under the specific conditions of Exp. I.

Lists. The nonsense syllable-adjective lists used by Underwood *et al.* (1962) were employed. The four, eight-pair lists, presented in the same order to all Ss, have the same stimuli and different responses, forming the A-B, A-C, A-D, A-E paradigm. Four orderings of the pairs were constructed for each list and were used equally often as starting orders.

Procedure. All lists were presented on a memory drum at a 2 : 2-sec rate. All Ss received the first three lists under MP (4-sec intertrial interval) for eight anticipation trials. List 4 was presented for six anticipation trials, half of the Ss learning by MP, half of the Ss learning by DP (3-min intertrial interval). The DP interval occurred between successive presentations of the list and was filled with symbol cancellation. Following the end of List-4 learning, Ss were reminded of their second session 48 hr later and asked not to serve in any other experiment between sessions.

On the paced recall (PR) test, Ss were asked to anticipate the correct List-4 responses. The drum turned at a 2 : 2-sec rate, but the response terms did not appear following the anticipation interval. The pairs were presented in the ordering which would have followed the last trial of List-4 learning.

The availability (AV) test was similar to the one employed by Keppel and Schwartz (1965) which consisted of a sheet of paper on which the eight nonsense syllables were listed in a column. Opposite each syllable were three blank spaces and the correct List-4 response. Fourth-list responses were provided to minimize differences in List-4 availability for the MP and DP Ss. The Ss were given 5 min to recall the responses from Lists 1–3. They were free to write down the words in any order and were not forced to indicate particular syllable-adjective pairings, although most Ss chose to do so. There were four AV tests, one corresponding to each of the four orderings of the List-4 pairs. The particular order each S received was decided by the same method employed for the PR test.

For the list-differentiation (LD) test, Ss were presented a listing of the syllables together with each of the four appropriately paired responses. Subjects were asked to indicate the order in which each of the four responses was learned. They were allowed unlimited time to complete the LD test and were forced to indicate list membership for every word. As with the other tests, there were four orderings of the syllables. The eight sets of four responses were each scrambled in a different ordering; the actual learning order was not used. In addition, two independent orderings of the response terms were employed equally.

Subjects. Each of the six experimental conditions contained 16 Ss. Subjects were either volunteers drawn from undergraduate classes in psychology at the University of California, in which service in experiments is a class requirement, or were paid students drawn from the University employment service. The experimental conditions were randomized in blocks of six, with starting and LD orders being assigned randomly under the restriction that equal numbers of each order be present. Subjects were assigned to the particular condition in order of appearance at the laboratory. While the numbers of volunteer and paid Ss were not equal, the assignment of these Ss was balanced over the six main experimental conditions. The 4 Ss who failed to return for recall were replaced immediately by the next S to report for the experiment.

Results

Lists 1–3: Learning. All Ss learned Lists 1–3 by MP. For all groups the average numbers of correct responses over the

eight trials were 35.97, 37.33, and 43.14, for Lists 1–3, respectively. None of the comparisons between MP and DP, nor among the recall treatments within these conditions, was significant (all p's > .25). These comparisons indicate a general comparability of groups in the learning of Lists 1–3.

List 4: Learning. The average performance of the combined MP and DP conditions for the six anticipation trials was 30.83 and 26.04, respectively, $F(1, 90) =$ 7.32, $p < .01$. None of the nested comparisons (within MP or DP) was significant (all p's > .10). The superiority of MP over DP is predicted by the extinction theory in that the DP interval should allow interfering associations to recover and slow down acquisition. One index of this recovery can be found in the incidence of interlist intrusions during List-4 learning. While the numbers are small, there were almost twice as many such intrusions for DP than for MP (26 vs. 15). These findings correspond to those obtained by Underwood *et al.* (1962).

Paced Recall (PR). The mean numbers of responses correctly recalled after 48 hr were 2.00 and .94 for DP and MP, respectively. While this difference approached significance following a Freeman-Tukey transformation, $F(1, 30) = 3.50$, $p < .10$, there was a reliably greater loss for MP than DP, $F(1, 30) = 6.46$, $p < .05$, when an adjustment was made for differences in degree of learning by

means of a probability analysis (Underwood, 1964). The number of interlist intrusions at recall was greater for MP than DP (42 vs. 32). The percentages of intrusions per overt emission (errors + correct responses) were 71.3% and 51.3%, respectively, but this difference was not reliable ($p > .05$). In short, the pattern of results observed on the PR test replicates the essential features of the earlier studies.

Availability (AV) Test. The extinction theory predicts greater loss of prior-list availability for the DP condition. Recall was measured in two ways, a stringent score referring to correctly paired responses and a lenient score counting all responses recalled, regardless of pairing. The mean numbers of correct recalls, for both measures, are presented for each list and condition in Table 1. For both measures, instead of inferior recall for the DP condition, recall is higher, but neither of these differences is significant (F's < 1). The only significant effect in the two analyses was for Lists, recall decreasing as a function of the number of interpolated lists, $F(2, 60) = 26.81$ and 22.89, $p < .01$. While it is true that this function is associated with a single ordering of the three lists, an unpublished analysis of a previous experiment (Keppel and Schwartz, 1965) showed these lists to be of equivalent difficulty ($F < 1$). Thus, the differences in recall as a function of list are probably not confounded by differences in list difficulty.

A final analysis of these data involved

TABLE 1

MEAN NUMBERS OF CORRECT RECALLS, SCORED TWO WAYS, FOR LISTS 1–3

Condition	Scoring	List 1	List 2	List 3	Total
MP	Stringent	1.38	1.44	3.62	6.44
	Lenient	1.44	1.75	3.75	6.94
DP	Stringent	1.19	2.25	3.56	7.00
	Lenient	1.56	2.62	3.62	7.81

TABLE 2

MEAN NUMBERS OF RESPONSES RECEIVING CORRECT LIST DESIGNATION

Condition	List 1	List 2	List 3	Total (Lists 1-3)	List 4
MP	5.56	4.12	3.31	13.00	3.94
DP	5.12	3.38	3.44	11.94	5.00

an inspection of lenient recall as a function of time of recall. It was conceivable, for example, that while MP and DP did not differ on the 5-min AV test, the two groups might differ in rate of prior-list recall. For this analysis the values for the three lists were combined and frequency of recall as a function of successive recall periods (first four 30-sec intervals and the remaining three 1-min intervals) was compared for the two conditions. The results of this analysis were negative, however, with both groups showing a negatively sloped exponential curve and little difference at any point during the recall period.

List Differentiation (LD). The mean numbers of correctly designated response members for the four lists and the two learning conditions are presented in Table 2. The values presented in Table 2 indicate a slight superiority of the MP Ss in discriminating Lists 1–3 and, as predicted by both the extinction and differentiation hypotheses, a large difference in the discrimination of List 4 in favor of the DP condition. This interaction is significant, $F(1, 90) = 9.28$, $p < .01$. An analysis of the simple effects shows no difference between MP and DP in the discrimination of Lists 1–3 ($p > .05$) and significantly better discrimination of List 4 for the DP condition, $F(1, 90) = 6.43$, $p < .025$. Finally, the progressive reduction in list differentiation from List 1 to List 3 was reliable, $F(2, 90) = 23.68$, $p < .01$.

An inspection of Tables 1 and 2 indicates that AV and LD scores represent opposite functions of list order, AV in-

creasing and LD decreasing from the first to the third list. This difference in trends was significant, $F(2, 120) = 47.84$, $p < .01$, and the two opposed functions were both reliable as simple effects (p's $< .01$). For Lists 1–3, then, these findings suggest that LD does not simply reflect differences in associative strength. It is conceivable that the underlying process is related to that responsible for serial-position phenomena, since a bowed LD curve is obtained for the four lists of the MP condition.

EXPERIMENT II

The results of Exp. I indicated that conditions which produce the DP retention effect are not associated with lowered prior-list availability. Similar findings were obtained following five relearning trials by Keppel (1964). It is possible, of course, that the AV test employed in Exp. I does not give a precise estimate of interlist interference at the time of recall. For example, the mere recall of the first three lists may eliminate differences in availability which may have been responsible for the DP effect. While the resolution of this issue must await the development of additional techniques by which the extent of interlist interference may be assessed, it is still reasonable to ask whether the DP effect is to be attributed to a reduction of interlist competition during recall. That is, if it can be shown that the DP effect need not be attributed to competition factors, then the search must be made for other explanations of the DP effect. The purpose of Exp. II was to provide an

estimate of competition. Such a measure has been suggested by Postman (1962) and consists of a comparison of paced and unpaced recall tests, the former being influenced by competition factors and the level of response availability and the latter being influenced by response availability alone. If the difference in recall on the paced and unpaced tests is compared with the level of recall on the unpaced test, a ratio is obtained which indexes the percentage of available responses which are lost through (a) the effect of competition and (b) the effect of pacing per se. If it is assumed that the noncompetitive aspects of the paced test are equivalent for MP and DP, any difference in the percentages will reflect the differential operation of competitive factors.

Method

Two groups of 15 Ss each were administered the MP and DP conditions outlined in Exp. I. On the recall test, 48 hr later, Ss were asked to write down List-4 responses opposite the appropriate stimuli. The responses from Lists 1–3 were provided, appropriately paired with their respective stimuli. Subjects were urged to guess and were allowed to list the words at the bottom of the page. A 4-min time limit was enforced. Subjects were drawn from the same populations described in Exp. I

Results

Learning. The two groups did not differ in the learning of the first three lists ($F < 1$). The mean numbers of correct responses over the three lists were nearly identical for Exps. I and II (116.46 and 114.40, respectively). As in Exp. I, MP was superior to DP over the six List-4 learning trials (29.75 vs. 24.21 correct responses), but this difference was not significant ($p < .25$); these values compare quite favorably with those obtained in Exp. I (30.83 and 26.04). Thus, it is reasonable to conclude that the MP and DP

groups of the two experiments were equivalent in learning ability.

Recall. The average numbers of correct responses for MP and DP, respectively, were 2.67 and 4.00, stringently scored, and 2.87 and 4.27, leniently scored. An analysis of loss scores (Trial 6 minus Recall) showed these MP/DP differences to be significant, $F(1, 28) = 4.22$ and 5.05, respectively, p's $< .05$. These results indicate that the DP effect is in part due to increased availability of List-4 responses at the time of recall, a finding which is not predicted by either the extinction or differentiation hypotheses.

In Exp. I the average paced recalls for MP and DP were .94 and 2.00, respectively. The differences between the paced and unpaced tests were 1.73 and 2.00; relative to the appropriate unpaced baselines, these values become 64.8% and 50.0%, respectively. If it is assumed that the noncompetitive effects of pacing are equivalent for the two conditions, this difference of approximately 15% indicates that DP has resulted in some reduction of competition on the paced test. Thus, the DP effect appears to be a function of two factors, competition and availability, but the factor of availability is the more important.

EXPERIMENT III

One final aspect of the unlearning-recovery theory which came under scrutiny was the question of the role of interlist interference. Underwood and Schulz (1961) did not obtain a DP effect with a single paired-associate list, but they did find the effect following a build-up of stimulus-specific interlist interference. While the effect does not appear with a single paired-associate list, where interlist interference is minimal and the DP intervals are short (e.g., 1 min), it is possible that the effect may occur when the DP

intervals are long (e.g., 24 hr). The purpose of Exp. III was to obtain estimates of the effect of widely spaced DP intervals on retention in the absence of specific interlist interference.

Method

Subjects were presented a single list of paired associates for eight anticipation trials at a 2 : 2-sec rate. Following the procedure of Keppel (1964), one half of the Ss learned under MP (4-sec intertrial interval) and one half learned under DP (24 hr separating every two anticipation trials). The list corresponded to the fourth list of the present experiments and the experiments of Keppel. The retention test consisted of five relearning trials taken either 1 or 8 days following the end of learning. The correct response was shown on all relearning trials. The Ss were introductory psychology students at Northwestern University, drawn from the same population sampled in the earlier study. There were 12 Ss in each of the four experimental groups. While the Ss were randomly assigned to the MP and DP groups within the two retention intervals, the two 1-day conditions were completed before the two 8-day conditions were begun.

Results

Learning. Since the learning curves for the two sets of MP and DP groups (one at each retention interval) failed to reveal any reliable differences (Fs < 1), these scores were combined. The performance of the combined MP and DP groups during learning is presented in Fig. 1, together with the corresponding data for the A-C paradigm (Keppel, 1964, Exp. I). The MP group represented here received all four lists during the same learning session. Although the two curves of the present experiment lie between the two curves of the four-list experiment, it is clear that the two single-list curves mirror the patterns obtained for the multiple-list experiment. The superiority of the multiple-list MP condition probably reflects the positive effects of learning-to-learn and warm-up. On the other hand, the inferiority of
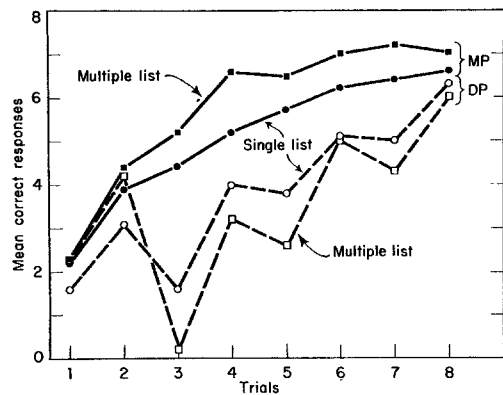

Fig. 1. Mean number of correct responses as a function of trials, MP and DP, and Experiments. The data for the multiple-list experiment come from Keppel (1964, Exp. I).

the corresponding DP condition is most likely due to the occurrence of strong interlist interference during learning.

Recall. The mean retention losses (Trial 8 minus Recall) for the various groups are plotted in Fig. 2. For the single-list conditions, DP resulted in a reduced retention loss at both intervals. The overall difference between MP and DP was significant, $F(1, 44) = 19.52, p < .01$, as was the main effect of Interval, $F(1, 44) = 16.78$, $p < .01$. The significant interaction of the two main effects, $F(1, 44) = 7.03, p < .025$, indicated that the difference between MP
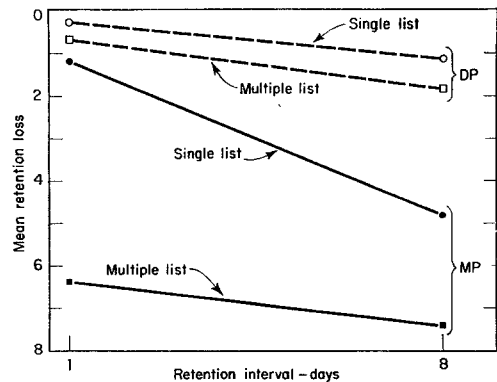

Fig. 2. Mean retention loss as a function of MP and DP and Experiments.

and DP was reliable for the 8-day interval only (p's > .10 and < .01 for 1 and 8 days, respectively). In terms of percentages, the DP Ss recalled 95.1% and the MP Ss recalled 80.8% after 1 day; corresponding values for the 8-day conditions were 81.4% and 28.8%.

A comparison of the single-list and multiple-list curves demonstrates the impressive facilitative effect produced by the 24-hr DP intervals, retention losses after 1 and 8 days being only slightly greater for the multiple-list condition. This small difference should be contrasted with the MP conditions, where the multiple-list group has essentially reached a point of maximum loss after one day. It is clear that a large portion of the DP effect observed by Keppel (1964) must be attributed to some process which allows the DP Ss to cope with the massive amount of interlist interference impinging upon the MP Ss at the time of recall.

DISCUSSION

Taken as a whole, these three experiments question the assumptions underlying the extinction-recovery theory. The facts that (a) prior-list availability is unaffected by DP (Exp. I and Keppel, 1964) and that (b) the DP effect is primarily due to increased availability, rather than the result of decreased interlist interference (Exp. II), are extremely damaging to the theory. In addition, the occurrence of a DP effect under conditions of minimal interlist interference (Exp. III), where DP consists of widely spaced intervals, limits the generality of the conclusion that the DP effect is largely dependent upon specific interlist relationships. The differentiation hypothesis, described earlier, did not fare much better in that it (a) does not predict a DP effect on an unpaced test (Exp. II), (b) does not account for the failure to obtain a DP effect with the A-Br paradigm (Underwood et al., 1962), and (c) does not predict the occurrence of the DP effect with a single paired-associate list (Exp. III). Furthermore, Houston (1966) was unable to show that a change in method per se was sufficient to produce the DP effect. As it stands, then, neither hypothesis is currently able to handle the facts which have been established in the present experiment and in the experiments reported previously. While both explanations were designed to account for the DP effect associated with short and long intervals, it is conceivable that the phenomena of the two time intervals reflect the operation of different basic processes. The explanation which will be offered here, however, will represent an attempt to deal with the two phenomena as the product of a single process.

For the sake of convenience, the discussion will center upon the findings of Exp. III, where widely spaced DP intervals produced superior retention of a single paired-associate list. The explanation will then be extended to the multiple-list situation and to the phenomena obtained with the shorter DP interval. First, it will be assumed that during acquisition S engages in a search for a linkage between some aspect of the stimulus and the required response. This association may be more or less direct or may involve a set of intermediate, i.e., mediational, linkages. It is further assumed that these associations will vary in strength and that Ss will vary in the success with which stable associations are established. Finally, it is assumed that the associations which are forgotten over the DP intervals will be the weakest and most labile of those which were formed during learning. This selection should have at least two consequences, (a) the stronger and more stable associations are selected, and (b) the selection process provides S

with a new opportunity to develop more satisfactory associations for the pairs which had been missed. The first point indicates that a strong association, one which is resistant to forgetting for whatever reason, will survive the DP interval and be maintained throughout the remaining acquisition trials. The assumption underlying the second point is that S, when confronted directly with the inadequacy of the previous association, attempts to form a new linkage or elaborates upon the original mnemonic device. Obviously, S will attempt to form this association during the study interval following recall. In addition, however, S may be able to utilize the time released during the exposure of the successful associations, which have survived the DP interval, to rehearse and perhaps to construct these new associations. Whether this renewed acquisition is visualized as consisting of a building upon of an old, but weak association or as a new sampling of potential mnemonics, is not important for this argument. What is critical is that S, by his failure, is given another attempt to replace an unstable association with a potentially stronger and more stable one.

In summary, if it is assumed that considerably less forgetting occurs over the relatively short MP intervals, this selection process will not operate to any great degree under these conditions. Thus, the MP S will end up with a collection of associations which represent the full range of associative strength, while the DP Ss, on the other hand, will have acquired a collection of relatively strong and stable associations. This difference in associative strength, then, accounts for the beneficial effects in retention of materials learned under DP. It should be noted that this explanation places the locus of the DP effect in the *acquisition* process, while the previous explanations have stressed perform-ance factors operating at the time of *recall*.

The present analysis may be applied easily to the multiple-list condition, by assuming the operation of the same type of selection process. In fact, the selection may be more efficient since the high interlist interference allows for a more severe and effective filtering of associative strength. The assumption of a selection process also allows for an accounting of the various phenomena associated with the relatively short DP intervals. These phenomena will be listed and discussed as a series of points: (a) The poorer critical-list learning under DP presumably represents the basic condition necessary for the operation of the selection process, i.e., S must fail during critical-list learning. (b) The selection hypothesis does not directly account for the greater number of interlist intrusions during critical-list learning for DP and at recall for MP. However, it appears that the DP effect is not dependent upon the specific occurrence of interlist intrusions during critical-list learning (Keppel, 1964). (c) The magnitude of the DP effect depends upon the length of the DP interval; this is true whether the variation involves the use of short intervals or the comparison of short and long intervals. The positive relation between the length of the DP interval and recall can be explained by assuming that the selection process is more efficient with the longer DP interval. (d) The fact that no DP effect is obtained in the absence of interlist interference (Underwood and Schulz, 1961) suggests that the selection process is dependent upon a certain degree of intertrial forgetting and that the amount of such forgetting represented by the single-list, short-interval situation is not sufficient to produce differences in selection and thus, retention.

The implication of the selection notion, of course, is that a DP effect should be

obtained with other transfer paradigms, provided there is sufficient opportunity for selection during learning. One obvious failure of this prediction is the absence of a DP effect with the A-Br paradigm when the DP interval is short. Considering the A-Br paradigm, where S is required to learn four successive pairings of the same stimulus and response terms, it is conceivable that S finds it relatively difficult to form stable associations. Therefore, the selection that is afforded by the short DP intervals is not sufficient to produce differences in associative strength which will be detected at recall. However, since a DP effect is found with the A-Br paradigm, when the DP intervals are long (Keppel, 1964), it is necessary to assume a greater opportunity for selection under these conditions and that this opportunity results in the necessary differences in associative strength to produce a DP effect.

A consideration of differences in item strength on the unpaced recall test (Exp. II) lends some support to the present analysis. A comparison of the strongest four and weakest four pairs in List-4 learning showed that the DP effect was largely associated with the stronger pairs. Specifically, the mean numbers of strong and weak responses recalled, respectively, were 1.36 and 1.30 for MP and 2.83 and 1.17 for DP. The interaction of Strength and MP/DP was significant, $F(1, 28) = 12.46$, $p < .01$. Since the selection notion holds that facilitation in recall is the result of selection following upon failure during learning, there should be some evidence of this selection and failure early in learning for the strong pairs. Because the first opportunity for selection occurred on Trial 1, strong and weak pairs were compared for MP and DP over Trials 1 and 2. These values appear in Table 3. The expectation of the theory is an early decrement for the strong DP pairs. Inspection of Table 3 reveals relatively poorer performance for the strong DP pairs on Trial 1, followed by a rapid increase on Trial 2. This observation is supported by a significant interaction of Trials × Strength × MP/DP, $F(1, 28) = 5.26$, $p < .05$. Thus, the pairs which seem to be associated with the difference in availability following

TABLE 3

MEAN NUMBERS OF CORRECT RESPONSES FOR STRONG AND WEAK PAIRS ON TRIALS 1 AND 2 FOR MP AND DP

	Strength			
	Strong		Weak	
Condition	Trial 1	Trial 2	Trial 1	Trial 2
MP	2.10	2.91	.30	1.02
DP	.83	2.17	.17	.30

DP in learning show some evidence for failure early during critical-list learning.

Admittedly, the evidence reported above is not extremely compelling by itself. On the other hand, the selection notion can account for most of the current findings which have been reported, while the hypotheses which have been offered earlier, extinction-recovery and differentiation, are unable to do so. Clearly, future experiments must be designed to test directly the selection hypothesis. One point of importance is the localization of the DP effect in the acquisition process rather than an identification with performance factors at recall. The unpaced recall test (Exp. II) offers a means by which the operation of performance factors may be neutralized, except, of course, if the unpaced test is not completely free from the effects of competition. Another critical consideration involves a specification of associative factors, i.e., forward, backward, and contextual associations (Keppel, 1964, Underwood, et al., 1962), necessary for the production of the DP effect. The selection hypothesis predicts a DP effect with other transfer paradigms, e.g., C-B and C-D, provided there is sufficient opportunity for selection of associative strength. At any rate, it is apparent that while there is no question as to the reliability of the DP retention effect, the theoretical analysis of this phenomenon is still not complete

REFERENCES

HOUSTON, J. P. List differentiation and distributed practice. *J. exp. Psychol.*, 1966, **72**, 477–478.

HOUSTON, J. P., AND REYNOLDS, J. H. First-list retention as a function of list differentiation and second-list massed and distributed practice. *J. exp. Psychol.*, 1965, **69**, 387–392.

KEPPEL, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. *J. verb. Learn. verb. Behav.*, 1964, **3**, 91–111.

KEPPEL, G., AND SCHWARTZ, B. Response availability and conditions of interpolated practice. *J. verb. Learn. verb. Behav.*, 1965, **4**, 489–493.

POSTMAN, L. Retention of first-list associations as a function of the conditions of transfer. *J. exp. Psychol.*, 1962, **64**, 380–387.

UNDERWOOD, B. J. Degree of learning and the measurement of forgetting. *J. verb. Learn. verb. Behav.*, 1964, **3**, 112–129.

UNDERWOOD, B. J., KEPPEL, G., AND SCHULZ, R. W. Studies of distributed practice: XXII. Some considerations which enhance retention. *J. exp. Psychol.*, 1962, **64**, 355–363.

UNDERWOOD, B. J., AND SCHULZ, R. W. Studies of distributed practice: XX. Sources of interference with differences in learning and retention. *J. exp. Psychol.*, 1961, **61**, 228–235.

(Received September 28, 1965)