

The Situation with Respect to the Spacing of Repetitions and Memory

ARTHUR W. MELTON¹

University of Michigan,² Ann Arbor, Michigan 48104

The revival of interest in the effectiveness of spaced practice, as compared with massed practice, in learning is attributed to the abandonment of the constraints of serial and paired-associate list learning and the discovery of stable benefits from spaced practice in continuous paired-associate learning, short-term memory for individual items, and single-trial free-recall learning. Comments are made about the preceding symposium papers by Underwood, Waugh, and Greeno, and some data on the differential effects of spacing of repetitions in free-recall learning are introduced in an effort to assess the current state of fact and theory.

Before looking at the data and theories that have been presented at this symposium, it seems to me worthwhile to consider how we have gotten to where we are today in the examination of the question of the relative effectiveness of massed practice (MP) and distributed practice (DP), and why.

I will not trace the history of research on MP vs DP in detail, but everyone knows that this issue has occupied experimental psychologists ever since the William James aphorism about learning to skate in summer, and swim in winter. Also, I suspect that most of you remember, at least vaguely, Jost's law, which dates from 1897, and states that "if two

associations are of equal strength but of different age, a new repetition has a greater value for the older one" (McGeoch, 1943, p. 140).

Over the last 30 years, we have witnessed intensive investigation of the MP-DP problem in a variety of behavioral contexts, but our ignorance is great and our understanding is to this day very limited—as attested by the search for viable hypotheses today. Being a confirmed optimist about the future of the science of human behavior, I would like to say why I think the focus of the symposium today—which is mostly concerned with understanding MP-DP effects in the free recall experiment—is a sign of progress.

¹ These comments are based on those made at the Midwestern Psychological Association symposium, May, 1969, but they have been liberally expanded and revised. In particular, it was necessary, in the interests of continuity and completeness, to insert substantial sections to cover what Tulving talked about at the symposium but would not write about. In the symposium I leaned heavily on his remarks because they were closely parallel to those I would have made if I had been a principal speaker. However, my statements about lag effects in free-recall learning, and my interpretation, are my responsibility alone, and are not intended to represent what he said.

² This work was supported by the Advanced Research Projects Agency, Department of Defense, and monitored by the Air Force Office of Scientific Research, under Contract No. AF(638)-1736 with the Human Performance Center, Department of Psychology, University of Michigan.

In the 1940's and decreasingly in the 1950's most of the effort on the MP-DP problem employed perceptual-motor skill tasks. The rotary pursuit task, other somewhat more analytic tracking tasks, the mirror-tracing task, and the inverted alphabet-printing task loomed large in our literature. Research interest has turned away from these tasks, especially during the last 10 years, not because the phenomena of reminiscence and MP vs DP were thoroughly understood, and a well-formed theory generated and tested, but because it was recognized that the basic theoretical issues could be examined in a more analytic fashion in simple associative learning

tasks involving verbal units, that is, in what has been quite superficially described as "rote learning."

Recognition of this truth began to occur in the early 1950's, as evidenced by some of our major learning theorists, especially Estes, turning to rote learning tasks to test and refine their theories. In doing this, there was a temporary abandonment, or postponement, of efforts to formulate exact theories of the relationship between repetition and learning, and similar fundamental issues, for trial-and-error learning, instrumental learning, and skill-learning tasks. In my opinion, this was because these latter tasks do not allow, as do rote learning tasks, the exact control over the frequency of occurrence of a specified event, nor the exactness of definition of the occurrence of a specified "correct" response, given a specified cue or stimulus.

In research on the MP-DP problem, the movement away from skill learning tasks and toward rote verbal-learning tasks mirrors this general trend in theoretical-experimental work on learning, and for very good reason. In skill-learning studies, repetition of the to-be-learned event was and is handled crudely by designating "trials" or "practice" in terms of time (usually seconds or minutes) that the Subject (*S*) is encouraged to interact with the task situation, and measurement of the change in performance from Trial *n* to Trial *n* + 1 is in terms of crude accomplishment measures, such as time-on-target, root-mean-square-error, or time to accomplish a fixed amount of work. Such controls of practice and measures of accomplishment are all very well as real-life definitions of practice and accomplishment, but are inappropriate for our theories, which must necessarily be about stimulus-response relationships or the organization of compounds of such relationships.

From the middle 1950's, and increasingly since then, attention has been given to the MP-DP problem in the context of verbal rote learning, harking back, in effect, to the flurry of interest and work by Jost in 1897, and to

the excitement about reminiscence in verbal learning in the 1930's. Unlike the skill learning tasks, the verbal learning tasks allow the specification that a particular verbal event, a word or trigram, has occurred for *x* seconds once, twice, or *n* times, or that a pair of such events has occurred once, twice, or *n* times. Such tasks also allow the identification of a required response event as having or not having occurred, with or without an appropriate retrieval cue.

These virtues of rote learning for analytic studies of the basic laws and properties of human learning and memory were, I suspect, always clear to those raised in the functionalist tradition of Ebbinghaus, Harvey Carr, John McGeoch, and Thorndike, and, therefore, to Underwood, Postman, and a few others who steadfastly turned to verbal learning for answers to questions about learning and memory. It is, therefore, significant for the point I wish to make next that Underwood embarked, in the 1950's, on a major effort to examine the MP-DP problem in the context of serial and paired-associate learning (Underwood, 1961). This effort was, for him, relatively disappointing, in large part because it was difficult to obtain large or consistent effects of distributed practice, as compared with massed practice, except when large amounts of response integration (as in the learning of Low-M CCCs as response terms) or large amounts of proactive inhibition were involved in the new learning.

Recent Positive Findings on Distributed Practice

The real renaissance of interest in MP vs DP as an issue of general theoretical importance arose from discovery of the marked effectiveness of DP on retention in verbal learning paradigms all of which depart from the traditional fixed-list rote-learning task that had been used with few exceptions by Underwood and all who preceded him. These "new" paradigms were the Continuous Paired-

Associate (CPA) method of Peterson, Saltzman, Hillner, and Land (1962), the Brown-Peterson method for investigating short-term memory for single verbal items (Peterson & Peterson, 1959), and the method of free-recall about which we have heard so much today.

Continuous Paired-Associate Learning

The first of these methods to show substantial differences in retention after MP and DP was the CPA method, as used in experiments by Peterson, Hillner, and Saltzman (1962) and Peterson, Wampler, Kirkpatrick, and Saltzman (1963). In this method the *S* is presented a list of paired words and numbers of indefinite length. Within this list a pair is presented one or more times before it is tested by the presentation of the stimulus term alone, and the number of other pairs (or tests of other pairs) that intervene between presentations and that intervene between the last presentation and test may be manipulated by the investigator without any constraints whatsoever, except for considerations of data-collection efficiency or the patience, or biological needs, of the *S*. Under these conditions, Peterson et al. and Young (1966) have shown that two spaced presentations of a pair improve the probability of recall of a response term, given the stimulus term, as compared with two massed presentations, if the retention interval after the last presentation is fairly long. This DP effect appears to be highly reproducible, in contrast to the earlier small and difficult-to-replicate effects of the massing or spacing of repetitions of lists of paired-associates. Furthermore, both Peterson et al. (1963) and Young (1966) have found that there is an orderly relation between the number of other pairs that occur between two presentations of a pair (the "lag" between presentations) and the amount of benefit from such spacing. The optimum lag is 7-8 intervening items when retention is measured 8-10 items after the second presentation, with retention declining with greater lags until a lag of 16 intervening items gives approxi-

mately the same retention as a lag of two intervening items.

Although strongly persuasive that a DP lag effect may be reliably obtained with the CPA method, it must be noted that the maximum DP effect so far obtained is about 25% improvement relative to the MP condition for two presentations, which will be seen to be a small effect compared to those found in free-recall learning. However, it must also be noted that we have been something less than avid in exploiting this discovery of DP effects in CPA learning, and have explored this DP lag effect with only a very limited subset of the possible paired-associate learning materials and the possible learning parameters with which such an effect might interact. Even so, it is, I believe, important that the DP effect has been found consistently in this paired-associate learning situation by the simple expedient of abandoning the list as a fixed unit and by manipulating the spacing of an individual S-R pair within a continuously changing context of other pairs. That is, the individual S-R association, which is what our theories are about, could have its spacing varied from a true zero, where presentations occur back-to-back, through any number of intervening different pairs (or tests of pairs), to a number of intervening pairs that would insure a zero probability of recall of the response, given the stimulus, at which time a second presentation of the pair could be introduced. Similar freedom is allowed in manipulating the duration of the retention interval before testing for the effects of the spacing of presentations.

Short-Term Memory for Individual Items

The second new memory method to reveal substantial beneficial effects of spaced presentations of to-be-remembered items is the Brown-Peterson method as employed by Peterson (1963) and Pollatsek (1969), whose results have been described by Greeno (1970). One important point about these data is that, as in the case of the CPA method, there is a strong "lag" effect, i.e., the beneficial effects of

spaced presentations increases as the interval between the two study presentations increases. That an inverted U-shaped function for lag has not been observed so far must certainly be attributed to our failure to establish conditions appropriate for its appearance. However, the data here, as in the case of the CPA method, provide straightforward evidence for what has been called the "strength paradox" (Bjork, 1969), which is that, over a certain range of lag values, the beneficial effect of a second presentation on later recall increases as the probability of recall at the time of that second presentation decreases. Finally, it is worthy of note that the beneficial effects of spaced presentations, as compared with massed presentations, are relatively small—even though reliable and orderly—in the Brown-Peterson method, as in the CPA method. The maximum beneficial effects so far observed are between 15% and 20%. As will be seen, this contrasts sharply with the large relative effects of spacing found in free-recall learning studies.³

Free-Recall Learning

The third experimental method that has revealed substantial and reliable effects of the spacing of presentations of items on memory for them is, of course, the free-recall learning method which has occupied the attention of Underwood (1970), Waugh (1970), and Tulving in this symposium. Despite the fact that the current flurry of research on the MP-DP problem with this method was sparked by Waugh's (1963) reported failure to find a DP effect, subsequent research has revealed very large, readily replicable, beneficial effects of spaced presentations of items on later free recall of them, at least under some conditions. Unlike the situation with respect to the data obtained by the CPA method and the Brown-Peterson method, the pretheoretic problem

³ Keppel (1964) has observed relatively large (ca. 50%) improved recall under DP, as compared with MP, under conditions involving large amounts of proactive interference and multiple-learning trials.

with the free-recall learning method is to understand why a DP effect is obtained in some studies and not in others. We must first examine the characteristics of the DP effect, when found, and the conditions under which such effects are found, or not found, before any useful comments can be made on the theoretical notions that have been expressed by Waugh, Underwood, or Greeno.

A straightforward integrative summary of what we know about MP vs DP effects in single-trial free recall is simply not possible at this time, even though we now have many reports of strong DP effects in such learning. This confusion is epitomized by the opposed conclusions of Underwood and Waugh regarding the Total-Time Law (TTL), which Underwood rejected on the basis of his data and Waugh accepted on the basis of her data. Their conclusions flow principally from studies that focus on frequency of presentation of to-be-remembered items under MP and DP conditions, and ignore (by randomizing) the possible effect of number of intervening items (lag) between successive presentations of an item. Another substantial series of studies (represented in the symposium by Tulving) employ presentation frequencies of only one or two and focus on the effect of lag between presentations under DP conditions. These studies consistently show a strong overall DP effect, but one that increases in a monotonic, negatively accelerated manner as the lag between presentations increases. The evidence on these two points of issue will be examined in the order stated.

Number of repetitions. Underwood's (1970) studies of the relation between frequency of presentation and recall under MP and DP schedules yield consistent results over a wide range of conditions. In his present report he shows that a DP schedule always produces better recall than MP and more so the greater the frequency of presentation. This holds for lists of sentences with mixed or unmixed MP and DP schedules (Experiments I-II), for medium-*M* nonsense syllables at 2- and 5-sec

rates of presentation with mixed MP and DP schedules (Experiment IV), and for mixed MP-DP lists of common nouns with children (Experiment III) and with or without irrelevant mental operations between each presentation of a word (Experiment V).

These findings are in nominal conflict with those of Waugh (1970), especially her Experiment II. There she found, in a study involving lists of words that were unmixed with respect to MP and DP schedules, differences between MP and DP of a curious but very orderly sort. MP facilitated recall, relative to DP, at frequencies of 1, 2, and 3, gave equal recall for a frequency of 4, and gave poorer recall for frequencies of 6 and 8. In another study at the same rate of presentation (1 sec/word) with lists involving mixed schedules of MP and DP frequencies of 2, there were no differences between MP and DP, although differences in favor of DP occurred when the words were presented at a 4-sec rate.

Waugh's data from her Experiment II show that the probability of recall of an item that occurs only once is strongly determined by whether the multiple-repetition items in the list are presented on MP or DP schedules. She quite properly concluded that her findings reflect strong tendencies to time-share covert rehearsal of low-frequency items at the expense of high-frequency items in the lists involving only MP schedules. Can Underwood's strong evidence for DP effects be explained away as an artifact of uncontrolled rehearsal time-sharing, since his studies with words and trigrams involved mixtures of MP and DP schedules? I think not, but it is necessary to go to studies that Underwood (1969, Experiment I and III) published earlier in order to get the evidence. In Expt. III in this earlier report, Underwood compared MP and DP in lists of words that had only MP schedules or only DP schedules, as in Waugh's Experiment II. Some *Ss* learned two lists involving frequencies of 1, 2, 3, or 4 with all multiple-presentation words on a MP schedule; other *Ss* learned the two lists when those words were presented

on random-lag DP schedules. The rate of presentation was one word each 5 sec, with each word presented auditorily twice during that period. The average probabilities of recall for the two lists after 1, 2, 3, and 4 presentations in the MP lists were .30, .29, .39, and .40, respectively; the comparable values for the DP lists were .29, .43, .55, and .57. The major point of interest here is that in unmixed lists, the point of equiprobable recall is at the frequency of 1, as it should be if there is no differential rehearsal strategy affecting the base rate of recall for the once-occurring items in the two kinds of lists. This is in sharp contrast to the fate of the once-occurring item in the MP and DP lists in Waugh's Experiment II. Also, it is clear that the frequency function in Underwood's experiment has a much steeper slope for the DP lists than for the MP lists, as is the case in Waugh's study, but is nonlinear and does not intercept at the origin. Finally, it may be noted that Underwood's data for recall after 1-4 MP in the unmixed-list study just cited are remarkably comparable to his data (1969, Experiment I) for MP frequencies in a parallel study involving lists of mixed MP and DP schedules. In this study the probabilities of recall after 1, 2, 3, and 4 massed presentations were .26, .32, .37, and .43. Thus, the average probability of recall under MP conditions in the unmixed list was .34, and the average probability of recall under MP conditions in the mixed list was also .34. Furthermore, the performance on the DP items in the mixed and unmixed lists were comparable except at the highest repetition frequency. In the mixed list, the probabilities of recall after frequencies of 1, 2, 3, and 4 were .26, .43, and .64, and .72, which are to be compared with .29, .43, .55, and .57 in the unmixed-list experiment.

From the above evidence I am forced to conclude that Waugh's (1970) findings in her Experiment II cannot be used to infer that Underwood's (1970) experiments with lists comprised of mixed MP and DP schedules were heavily contaminated by uncontrolled

rehearsal that was differential for the MP and DP items. It is particularly important that the comparability of recall probabilities after 1–4 massed presentations in mixed and unmixed lists denies the hypothesis that performance under MP conditions is depressed by their being in a context of items on a DP schedule. These statements do not, of course, deny the validity of Waugh's findings, which are among the most orderly data available to us and must somehow be understood. My only suggestion on this score is a rather radical one, and perhaps foolish. It is that at fast rates of presentation of words for free-recall, *Ss* are more likely to adopt differential rehearsal strategies for MP items. Whether this or some better explanation of the discrepancy between Underwoods' and Waugh's findings is applicable, it is obviously of immediate importance that parametric experiments be made to determine recall after frequencies of at least 1–4 presentations, in MP and DP schedules, in mixed and unmixed lists, and with several rates of presentation ranging from one allowing only marginally adequate processing of each word (e.g., 2/sec), through an obviously adequate but fast rate (e.g., 1/sec), and through at least two slow rates that allow for different amounts of nominally uncontrolled rehearsal time (e.g., 1 word/2 sec and 1 word/4 sec). For the present, it seems clear that the Total-Time Law is in deep trouble as an empirical law, and it may or may not be salvageable as a theoretical law relating to the duration of *S*-determined, as contrasted with *E*-determined, processing time for individual items, as suggested by Waugh (1970).

Spacing of repetitions (lag). The deep trouble of the Total-Time Law, whether the empirical law or the theoretical law, becomes deeper if it can be demonstrated that the lag between presentations of an item under DP schedules is a variable of importance in free-recall learning. As mentioned earlier, neither Waugh (1970) nor Underwood (1970) accept lag as important. Underwood's rejection of lag is on the basis of an experiment, similar to those he reports in

the symposium, in which both frequency (2, 3, and 4) and number of words between repetitions of a word (2, 8, 14, and 20) were varied within a long list (Underwood, 1969, Experiment IV). This experiment showed a significant effect of frequency, lag, and frequency \times lag interaction, but no orderly statement of the interaction could be formulated. It will require replication before it may be accepted as evidence. Waugh (1970) on the other hand, rejects the notion of a lag effect on the basis of evidence obtained from essentially the same experimental design (her Experiment I, this symposium) that has repeatedly yielded a lag effect in the studies reported by others.

The essential feature of the design is that a word is given only one or two presentations; if given two presentations, these are massed or spaced by inserting different numbers of intervening words (from 1 to as many as 40 in different experiments). In the studies to be referenced all massing and spacing conditions for a word occur within the list that is learned (i.e., it is a "mixed" list), and they are preceded by a "primacy buffer" of 5–8 items which occur only once and are followed by a "recency buffer" of 5–8 items which occur only once.

In 1963 Waugh summarized several experiments on the lag effect in free recall which produced essentially the same findings she reports in this symposium (Experiment I) for the 1-sec rate, and it is notable that in this more recent experiment she obtained a DP effect, but no lag effect, when she used a 4-sec rate of auditory presentation. To my knowledge, these are the only negative findings on the effect of lag with this experimental design. After a brief preliminary report of increasingly beneficial effects of lags of 0, 8, 20, and 40 in the free recall of words (Melton, Reicher, & Shulman, 1966), Melton and Shulman (1967) reported the data shown in Figure 1. In this experiment, each *S* had preliminary practice on free recall of two-digit numbers which were presented under MP and different-lag DP conditions and then was given one-trial recall tests on three lists of 48 different four-letter nouns. In the

middle of each list were eight words that occurred once and four words that occurred twice at lags of 0, 2, 4, 8, 20, and 40. Different groups of 48 paid college students learned

these lists by visual presentation at the rate of 1.3, 2.3, or 4.3 sec per word (of which .8 sec was dark time involved in changing slides). It is clear from Figure 1 that the main effects

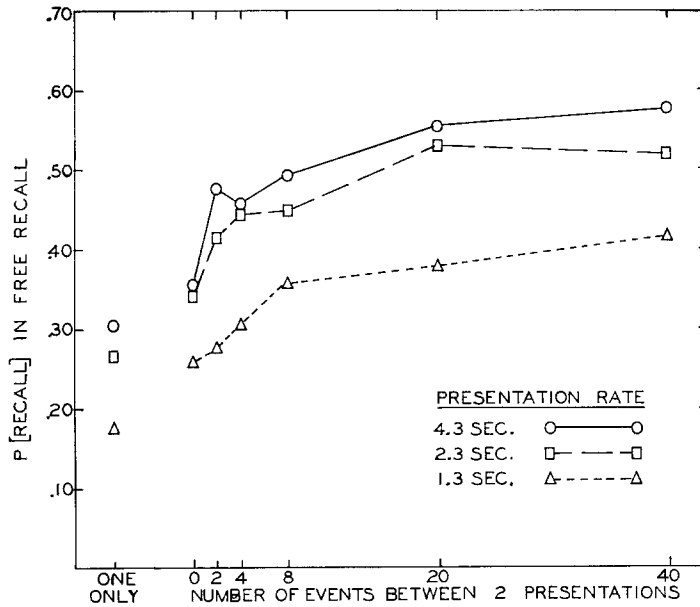


FIG. 1. Probability of recall of words that occur once or twice, with varying numbers of events between the two presentations, when presented visually at rates of 1.3, 2.3, and 4.3 sec/word. (Melton & Shulman, 1967.)

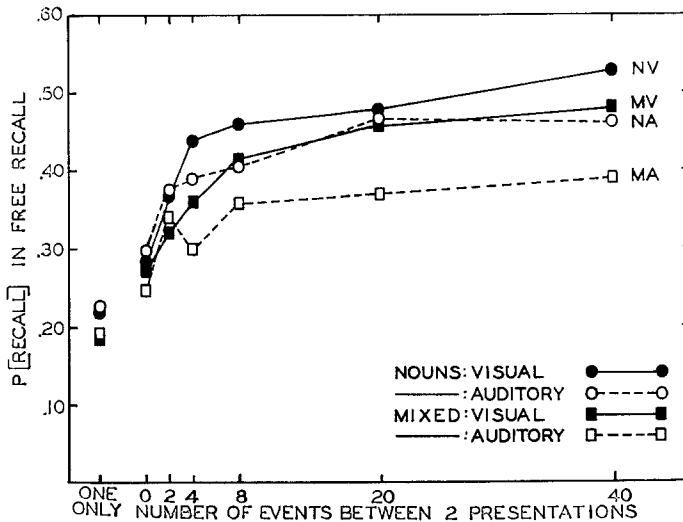


FIG. 2. Probability of recall of high-frequency four-letter nouns and words of mixed frequency, word-class, and length when these words occur once or twice, with varying numbers of events between the two presentations, after visual and auditory presentation at the rate of 2.3 sec/word. (Data from Melton & R. A. S. Adams, unpublished.)

of rate of presentation and lag were significant and that there was no interaction of rate and lag. While it may be, as Waugh (1970) says, that rate of presentation is a critical variable in determining the DP effect when presentation is auditory, it is clearly not a critical variable when presentation is visual.

This experiment was followed by a very similar one (not previously reported) by Melton and R. A. S. Adams in which each of 192 college *Ss* learned four lists at the 2.3 rate of presentation. Two of the lists were made from high-frequency four-letter nouns (as in the experiment on rate), and two lists were made from mixtures of high- and low-frequency words of different word classes ("mixed" words) of the sort Waugh (1963) used in her original study. Each *S* had one list of each word type with visual presentation and one with auditory presentation, with complete counterbalancing of lists and words-within-conditions across *Ss*. The outcome is shown in Figure 2. The main effects of mode of presentation (better recall with visual), type of word (better recall with homogeneous nouns), and lag were significant, and the modality \times lag interaction was significant. While all four curves show an effect of lag, it is clear that the slope of the lag function is greater for visual presentation than for auditory presentation, and that auditory presentation with the mixed words produces a lag slope that is small indeed.

While it is tempting to conclude from these last findings that we would have replicated Waugh's (1970) failure to obtain a lag effect if we had presented our words auditorily at 1/sec, I am not yet prepared for this conclusion. Other possibly important differences between her experiments and ours are that she used quite short lists (33 words in her Experiment I), a fixed and readily perceived pattern of occurrence of her lag exemplars in all lists, and many lists per *S*. Of more importance at this time is the generality of the finding of a strong effect of lag under DP schedules. In addition to the studies cited above, our own

unpublished studies have obtained significant effects of lag in the free recall of high-*M* and low-*M* CVCs (thus confirming and extending Underwood's Experiment IV), and we are currently finding that the lag effect can be accentuated by instructions regarding word-word encoding (subjective organization) and minimized by a variety of operations designed to interfere with word-word encoding. There are also the confirmatory findings of Madigan (1969) and his demonstration that the slope of the lag function can be reduced if *Ss* are presented with nouns with different associative modifiers and recall is cued.

As in the case of the effect of frequency of presentation under MP and DP schedules there remain many questions about the lag effect that will surely require systematic manipulation of procedural variables before they can be answered. So far there has been no indication that the lag between two presentations of an item can become so long that there is a decline in probability of recall toward that obtained with only one presentation. However, there is the strong suggestion that the effect asymptotes at or between 20 and 40 intervening words. The one thing that is obvious about the lag effect is that it is a very orderly relation, whenever it occurs.

Theoretical Issues

If one accepts the validity of the lag effect in free-recall learning it can then be said that comparable beneficial effects of DP and of the degree of spacing of distributed presentations have been observed in continuous paired-associate learning, in short-term memory for individual items, and in single-trial free recall. It would, of course, be a mistake to assume that the underlying causes of the DP effect in these three situations are the same merely because the spacing of presentations has what appears to be the same effect, but this circumstance certainly favors experimental-theoretical convergence on the necessary and sufficient conditions for beneficial effects of spaced repetitions. As Greeno (1970) has aptly noted,

the finding of a DP effect and an effect of lag in short-term memory serves as a corrective to several hypotheses that might account for some or all of the effect in CPA learning and free-recall learning. Similarly, the relative magnitude of the maximum DP effect and the functions relating lag to the size of the effect in these different reference experiments may well suggest hypotheses that relate the underlying causes of the DP effect to the characteristics of the learning task, which is what will be necessary before our current rediscovery of DP effects can become useful in managing learning processes. On the basis of the data currently available, it appears that the DP effect is very much greater, and is effective over very much longer lags, in free-recall learning than in the CPA or short-term memory experiments. Once again, however, we sorely need systematic data-gathering before we can hope for confirmation or disconfirmation of general theoretic interpretations. I suspect that we will also need ingeniously devised "transition experiments" (Underwood, 1964) that bridge the procedural gaps between CPA, short-term memory, and free-recall experiments as they are presently employed.

Even though any decisions about theoretic issues, or even limited hypotheses, must await more information from experiments, a few comments about general and specific interpretations may serve some useful purpose. In the first place, I find it comforting that no one in the symposium appealed to "consolidation" as an explanation of observed DP effects, but all sought instead to find understanding from hypotheses about the information-processing activities of the *Ss*. My bias against an appeal to consolidation is not because some such autochthonous process cannot be involved. It is because I see the formulation and testing of interpretations in terms of the *S*'s information-processing activities as more likely to expose the psychological factors underlying the relative effectiveness or ineffectiveness of DP under different conditions. There are as yet insufficient independently defined parametric

guidelines in consolidation theory for useful employment of the theory in the situations under discussion.

My second general comment relates to the class of hypotheses that attribute the presence or absence of DP effects to either the attention and effort given to storing an item under MP conditions or the *S*'s strategy to differentially rehearse MP and DP items. These things happen, but the question of interest is whether they explain the results of the MP-DP experiments. Underwood (1970) worried that his strong evidence for beneficial effects of DP may merely reflect a failure of attention and learning effort under MP conditions; Greeno (1970) decided that *S* must merely "turn off (or turn down) the processor that transfers items to long-memory"; Waugh (1970) accepted Greeno's hypothesis and stresses in addition the notion that any "free" time is used to process other items in the list. All three, therefore, believe or are suspicious that such factors explain the observed MP-DP effects.

It seems to me unlikely that these hypotheses explain the data presently available. As for the notion that *S* turns off or attenuates the processing of a word under MP conditions, it is not clear from Greeno's statement how this would account for the lag effect observed in the short-term memory experiments. Even though the spacing intervals used so far are rather limited in duration, they are beyond the range of primary memory involvement. It is even less clear how the notion would account for the increasing recall with increasing interpresentation lag in the free-recall experiments where a lag of 20 intervening items regularly produces better recall than a lag of eight intervening items. Eight intervening items is usually considered sufficient to clear the primary memory buffer of an old item, and should be sufficient to restore full processing activity.

The insufficiency of the hypothesis that *Ss* adaptively and differentially time-share their rehearsal of the presented item and other items

is more difficult to prove because it appears to be a pervasive intellectual skill, at least among college students. Waugh (1970) provided very persuasive evidence for its importance in her experiment. As previously suggested, such skills may be uniquely involved in her Experiment II because of the time-pressures of fast presentation. On the other hand, Greeno's (1970) rejection of the rehearsal time-sharing hypothesis in his paired-associate experiments on the basis of Potts (1969) data may be unwarranted. Greeno relinquished his version of the rehearsal time-sharing hypothesis because Potts failed to find evidence that the pairs immediately before the MP pairs were recalled better than those farther back in the sequence. It is possible that the expected recency gradient for the benefits of rehearsal during "free" time on the DP pair was eliminated or damped because the task of *S* was to categorize stimulus words as belonging to response classes 1, 2, 3, 4, and 5. Under these circumstances, *S* may have imposed a selective rehearsal on recent pairs that had the same response terms as the now-presented pair. If the five pairs prior to the DP pair had the same response term as the DP pair with equal probability, as would be the case in a well-designed experiment, this selective rehearsal strategy of *S* would eliminate a recency gradient.

Acceptance of the evidence that *Ss* do adaptively and differentially time-share their rehearsal under MP and DP conditions, and differentially in mixed and unmixed MP and DP schedules, does not require acceptance of the proposition that these skills explain the observed DP effects. The main argument against such a proposition at this time is, in the case of free-recall learning, the orderly effect of lag on probability of recall. Any rehearsal time-sharing skill that is invented as an intervening variable to account for the observed lag effects will merely endow *Ss* with skills that mirror the observed relations. The previously suggested systematic studies of mixed and unmixed MP and DP schedules at different rates of presentation should pro-

vide more specific bases for evaluating such a notion. Meanwhile, it seems quite clear, as Greeno (1970) pointed out, that such rehearsal time-sharing is inappropriate as an explanation of the observed DP effects in short-term memory studies.

What other hypotheses are available in our effort to understand the observed DP effects? Tulving's paper in the symposium developed an alternative hypothesis that has guided my own work and the work of Madigan (1969). This hypothesis is that DP permits more different cues to be stored than does MP, and that these additional cues aid retrieval. It has been widely observed that normal free-recall learning involves subjective organization of word-word combinations and that these subjective units of two or more words serve as cuing systems at the time of recall (Tulving, 1968). The prediction of the beneficial effects of DP, and the gradation of this beneficial effect as a function of lag, is obtained by assuming, first, that the coding of a word in two different subjective units (or in a larger subjective unit) increases the cues or access routes to its retrieval, and second, that as the lag between two or more occurrences of a word increases, the word contexts in which it occurs become less and less correlated (more independent) and the total number of different cues to its retrieval increases.

It would be premature and inappropriate to attempt a detailed defense of this alternative hypothesis at this time. However, it is appropriate to point out that this hypothesis is attractive in part because it conceives the DP effect as an outcome of fundamental characteristics of man's processing of information into memory and retrieval of information from memory, rather than as an outcome of failure to control his effort, attention, or rehearsal strategies. In particular, it can be linked readily with the concept of stimulus encoding variability (Martin, 1968) and to the involvement of context in such encoding variability. In any event, the hypothesis will undoubtedly stimulate experimentation and much controversy—

which is the only way we learn new truths in science.

In summary, it may be said with confidence that the effects of distributed practice on the remembering of verbal items and the relations between them are available in sufficient magnitude, with sufficient replicability, and with sufficient variability in a variety of experimental situations, to warrant intensive systematic investigation and intensive theoretic efforts. It is possible that now, at long last, we are on the verge of understanding why and under what conditions repetition improves remembering. At least, we may have cleared the path to understanding by finding the Total-Time Law, and perhaps also a simple cumulative strength interpretation of the effects of frequency, to be valid only under very special circumstances, if at all.

REFERENCES

- BJORK, R. A. Repetition and rehearsal mechanisms in models for short-term memory. Ann Arbor: Human Performance Center Technical Report No. 14, 1969, University of Michigan.
- GREENO, J. G. Conservation of information-processing capacity in paired-associate memorizing. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 581-586.
- KEPPEL, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 91-111.
- MADIGAN, S. W. Intraserial repetition and coding processes in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 828-835.
- MARTIN, E. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 1968, 75, 421-441.
- MCGEOCH, J. A. *The psychology of human learning*. New York: Longmans Green, 1943.
- MELTON, A. W., REICHER, G. M., & SHULMAN, H. G. A distributed practice effect on probability of recall in free recall of words. Paper presented at Psychonomic Society, St. Louis, October, 1966.
- MELTON, A. W., & SHULMAN, H. G. Further studies of a distributed practice effect on probability of recall in free recall. Paper presented at Psychonomic Society, Chicago, October, 1967.
- PETERSON, L. R. Immediate memory: data and theory. In C. N. Cofer & B. S. Musgrave (Eds.) *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963. Pp. 336-353.
- PETERSON, L. R., HILLNER, K., & SALTZMAN, D. Supplementary report: time between pairings and short-term memory. *Journal of Experimental Psychology*, 1962, 64, 550-551.
- PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, 58, 193-198.
- PETERSON, L. R., SALTZMAN, D., HILLNER, K., & LAND, V. Recency and frequency in paired-associate learning. *Journal of Experimental Psychology*, 1962, 63, 396-403.
- PETERSON, L. R., WAMPLER, R., KIRKPATRICK, M., & SALTZMAN, D. Effect of spacing presentations on retention of a paired associate over short intervals. *Journal of Experimental Psychology*, 1963, 66, 206-209.
- POLLATSEK, A. W. Rehearsal, interference, and spacing of practice in short-term memory. Ann Arbor: Human Performance Center Technical Report No. 16, 1969, University of Michigan.
- POTTS, G. R. Distance from a massed double-presentation as a factor in paired-associate learning. Paper presented at meetings of the Midwestern Psychological Association, Chicago, May, 1969.
- TULVING, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1968. Pp. 2-36.
- UNDERWOOD, B. J. Ten years of massed practice on distributed practice. *Psychological Review*, 1961, 68, 229-247.
- UNDERWOOD, B. J. The representativeness of rote learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 47-78.
- UNDERWOOD, B. J. Some correlates of item repetition in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 83-94.
- UNDERWOOD, B. J. A breakdown of the Total-Time Law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 573-580.
- WAUGH, N. C. Immediate memory as a function of repetition. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 107-112.
- WAUGH, N. C. On the effective duration of a repeated word. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 587-595.
- YOUNG, J. L. Effects of intervals between reinforcement and test trials in paired-associate learning. Institute of Mathematical Studies in the Social Sciences Technical Report No. 101, 1966, Stanford University.

(Received June 16, 1970).