

# PSYCHOLOGICAL REVIEW

---

## PRIMARY MEMORY<sup>1</sup>

NANCY C. WAUGH

*Harvard Medical School*

AND DONALD A. NORMAN

*Center for Cognitive Studies, Harvard University*

A model for short-term memory is described and evaluated. A variety of experimental data are shown to be consistent with the following statements. (a) Unrehearsed verbal stimuli tend to be quickly forgotten because they are interfered with by later items in a series and not because their traces decay in time. (b) Rehearsal may transfer an item from a very limited primary memory store to a larger and more stable secondary store. (c) A recently perceived item may be retained in both stores at the same time. The properties of these 2 independent memory systems can be separated by experimental and analytical methods.

It is a well-established fact that the longest series of unrelated digits, letters, or words that a person can recall verbatim after one presentation seldom exceeds 10 items. It is also true, however, that one can nearly always recall the most recent item in a series, no matter how long the series—but only if this item may be recalled immediately, or if it may be rehearsed during the interval between its presentation and recall. Otherwise it is very likely to be lost. If we may assume that attending to a current item precludes reviewing a prior one, we can say that

the span of immediate memory must be limited in large part by our inability to rehearse, and hence retain, the early items in a sequence while attempting to store the later ones. Our limited memory span would then be but one manifestation of our general inability to think about two things at the same time.

Why should an unrehearsed item in a list be forgotten so swiftly? Is its physiological trace in some sense written over by the traces of the items that follow it? Or does this trace simply decay within a brief interval, regardless of how that interval is filled? Tradition, in the guise of interference theory, favors the first explanation (McGeoch, 1932; Postman, 1961), although some psychologists now think that new memory traces must fade autonomously in time (Brown, 1958; Conrad, 1957; Hebb, 1949). Until now, no one has reported any data

<sup>1</sup> This research was supported in part by Research Grants No. MH 05120-02 and MH 08119-01 from the National Institutes of Health, United States Public Health Service, to Harvard University, Center for Cognitive Studies and to Harvard Medical School, respectively. The second author was supported by a National Science Foundation Postdoctoral Fellowship at the Center for Cognitive Studies.

which clearly contradict either of these ideas. In fact, when we first considered the problem of the instability of recent memory traces, we thought it entirely possible that both decay and interference operate over brief retention intervals to produce forgetting, and we therefore designed an experiment to weigh their respective effects. The results of this experiment were unexpectedly straightforward—and seemingly inconsistent with certain other existing data on immediate retention. We have been able, however, to formulate a simple quantitative model which relates our results to those reported by other investigators. What began as an attempt to evaluate two very general hypotheses about the forgetting of recent events has therefore resulted in a specific theory of short-term memory.

We shall describe our experiment in Section I below. A major portion of this paper, Section II, will be concerned with the description and application of our model. In Section III we shall discuss this model in relation to the general question of whether short- and long-term retention represent distinguishably different psychological processes.

## I. PROBE-DIGIT EXPERIMENT

Our experiment was designed to measure the recall of a minimally rehearsed verbal item as a joint function of the number of seconds and the number of other items following its presentation. The general procedure was as follows. Lists of 16 single digits were prepared with the aid of a standard table of random numbers, under the constraint that no digit should appear more than twice in a row. The last digit in every list was one that had occurred exactly once before, in Position 3, 5, 7, 9, 10, 11, 12, 13, or 14. On its second appearance, this "probe-digit" was a cue for the recall of the digit that had followed it initially.

The lists were recorded on two magnetic tapes; they were read in a monotone voice by a male speaker at a constant rate of

either one or four digits per second. Each of the nine possible probe-digit positions was tested 10 times. The two tapes accordingly contained 90 test lists (plus 8 practice lists) apiece, all read at the same rate. The last digit in every list, the probe-digit, was accompanied by a high-frequency tone to aid the subject in detecting the end of the list. The position of the initial presentation of the probe varied randomly from list to list on each of the two tapes.

The subject's task was to write down the digit that had followed the probe digit in the list, guessing if he did not know. Since the probe-digit was unique in Positions 1 through 15, there was only one possible correct answer on any trial. Every subject listened to the list through earphones for a total of 12 experimental sessions, 6 with each tape, alternating between fast and slow lists. The first session under each condition and the first eight lists listened to in each session were considered to be practice and, unknown to the subject, were not scored.

The subjects received explicit instructions to control rehearsal by "thinking only of the last digit you have heard and never of any of the earlier ones." These instructions were repeated before the second session, and occasional reminders were given throughout the course of the experiment. Thus, the subjects were to rehearse every item during the interitem interval immediately following it. Our instructions were not designed to eliminate the rehearsal of single items as such, but rather to eliminate the rehearsal of *groups* of digits. The experiment actually tested the retention of a digit pair, the probe-digit and its successor. The retention of this pair should be independent of the interitem interval, if the instructions to avoid grouping were followed faithfully. We hoped, in effect, to test the retention of unrehearsed pairs of digits under two rates of presentation.

The subjects were four Harvard undergraduates, three males and one female.

The responses were scored and analyzed to yield a serial position curve for each rate of presentation, relating the relative frequency of an item's correct recall to its distance from the end of the list. A comparison of the two functions allows us to assess the relative effects of decay and interference on short-term forgetting, according to the following line of reasoning. Consider the recall of Item  $i$  from the end of the line. If the list was read at the rate of one item per second, then  $i$  items would have intervened, and  $i$  seconds would have elapsed between

the time the subject heard the item and the time he attempted to recall it. (We count the second appearance of the probe-digit as an intervening event.) If the items were read at the rate of four per second, on the other hand, then only  $i/4$ , rather than  $i$ , seconds would have elapsed between the occurrence of Item  $i$  and the subject's attempt to recall it. A total of  $i$  other items would, of course, still have intervened between these two events. Therefore, if the probability of recalling Item  $i$  from the end of a slow list were identical with the probability of recalling Item  $4i$  from a fast list, we could conclude that recent memory traces decay in time, independently of one another. Conversely, if the probability of recalling Item  $i$  were invariant with rate of presentation, we could conclude that rapid forgetting is caused primarily by retroactive interference.

The results of the experiment are shown in Figure 1. The scores for the individual subjects are presented in Figures 1A and 1B. The pooled data, corrected for guessing, are shown in Figure 1C.<sup>2</sup> Each point in Figures 1A and 1B is based on 50 observations; each point in Figure 1C, on 200. It is evident that there are consistent differences among subjects, but little interaction between subjects and serial positions. Furthermore, although there appears to be a slight interaction between relative frequency of recall, or  $R(i)$ , and rate of presentation, it is clear that the effect of rate is relatively small compared to the effect of serial position. The main source of forgetting in our experiment was interference.

The differences between the two sets of points shown in Figure 1C are not statistically reliable, according to an analysis of variance performed on the number of items recalled by each sub-

<sup>2</sup> The response set—the 10 digits—was known to the subjects, and they knew that the probe would not be the same as the test digit. Thus the probability of correctly guessing the answer,  $g$ , was  $1/9$ . A standard normalizing technique was used to eliminate the effects of guessing from the data, namely,  $\hat{p}(\text{recall}) = [p(\text{correct}) - g]/(1 - g)$ .

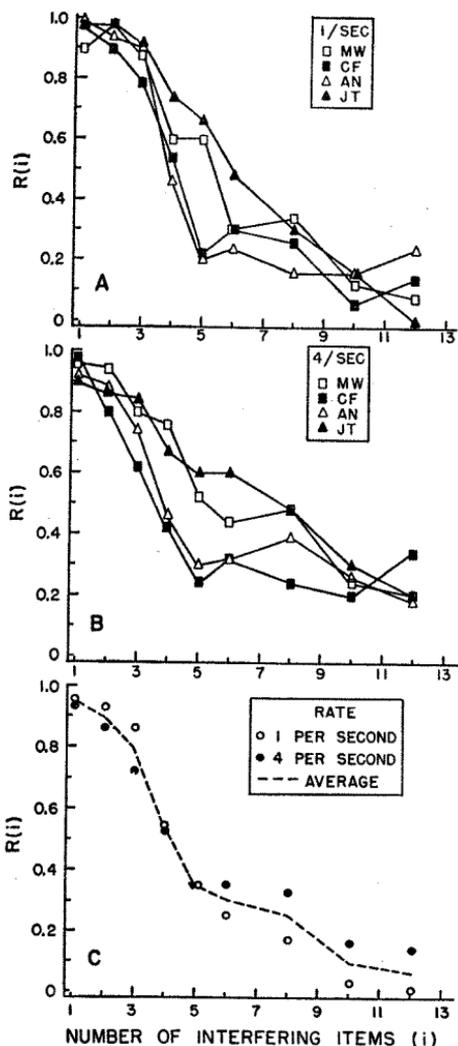


FIG. 1. Results of the probe-digit experiment. (Figures 1A and 1B represent retention functions for individual subjects under two rates of presentation; in Figure 1C these data have been pooled.)

ject at each value of  $i$  under the two rates of presentation ( $F < 1$  for the mean square between rates tested against the interaction between subjects and rates). This conclusion is borne out by the results of nine Kolmogorov-Smirnov two-sample tests, one for each value of  $i$ , performed on the distributions of number of items recalled per subject per session under

the two rates of presentation. We have therefore fitted the points shown in Figure 1C with a function that represents the probability of recalling Item  $i$  from the end of a series, estimated across rates of presentation. This function decreases monotonically with  $i$ , attaining a value of about .07 at  $i = 12$ .

## II. MODEL FOR PRIMARY MEMORY

When we compared the foregoing results with the typical outcome of the first trial in a standard list-learning experiment, we found ourselves facing two dilemmas. In the first place, it often happens that an item in a long list is recalled after 10 or 20, or even more, items have followed it. But in our experiment, probability of recall was effectively zero for the eleventh item in from the end of a list. In the second place, various investigators have shown that probability of recall increases with presentation time (see Posner, 1963), yet in our experiment this probability, for all practical purposes, was independent of the rate at which the digits were read.

In seeking for a way to account for these discrepancies, it occurred to us that one difference between our experiment and previous ones in this area is that we instructed our subjects not to think about any item in a list once the next had been presented. This instruction to avoid rehearsal is, to be sure, rather unorthodox, although not completely without precedent (Underwood & Keppel, 1962). In order to minimize rehearsal, many experimenters try to keep the subject so busy that he does not have time to rehearse; but we think it highly likely that a well-motivated subject who is trying to learn a list will rehearse unless specifically enjoined from doing so. The typical subject's account of how he learns a list (Bugelski, 1962; Clark, Lansford, &

Dallenbach, 1960) bears us out on this point. In fact, it is probably very difficult *not* to rehearse material that one is trying to memorize.

We shall assume here that rehearsal simply denotes the recall of a verbal item—either immediate or delayed, silent or overt, deliberate or involuntary. The initial perception of a stimulus probably must also qualify as a rehearsal. Obviously a very conspicuous item or one that relates easily to what we have already learned can be retained with a minimum of conscious effort. We assume that relatively homogeneous or unfamiliar material must, on the other hand, be deliberately rehearsed if it is to be retained. Actually, we shall not be concerned here with the exact role of rehearsal in the memorization process. We are simply noting that, in the usual verbal-learning experiment, the likelihood that an item in a homogeneous list will be recalled tends to increase with the amount of time available for its rehearsal. The probe-digit experiment has shown, conversely, that material which is not rehearsed is rapidly lost, regardless of the rate at which it is presented. It is as though rehearsal transferred a recently perceived verbal item from one memory store of very limited capacity to another more commodious store from which it can be retrieved at a much later time.

We shall follow James (1890) in using the terms *primary* and *secondary memory* (PM and SM) to denote the two stores. James defined these terms introspectively: an event in PM has never left consciousness and is part of the psychological present, while an event recalled from SM has been absent from consciousness and belongs to the psychological past. PM is a faithful record of events just perceived; SM is full of gaps and distortions. James believed that PM extends over a fixed

period of time. We propose instead that it encompasses a certain number of events regardless of the time they take to occur. Our goal is to distinguish operationally between PM and SM on the basis of the model that we shall now describe.

Consider the general scheme illustrated in Figure 2. Every verbal item that is attended to enters PM. As we have seen, the capacity of this system is sharply limited. New items displace old ones; displaced items are permanently lost. When an item is rehearsed, however, it remains in PM, and it may enter into SM. We should like to assume, for the sake of simplicity, that the probability of its entering SM is independent of its position in a series and of the time at which it is rehearsed. Thus, it would not matter whether the item was rehearsed immediately on entering PM or several seconds later: as long as it was in PM, it would make the transition into SM with fixed probability. (Our PM is similar to Broadbent's, 1958, *P* system. One difference between our two systems is that ours relates rehearsal to longer term storage, whereas his does not.)

Finally, we shall assume that response-produced interference has the same effect on an item in PM as does stimulus-produced interference. That is, the probability that an item in PM will be recalled depends upon (a) how many new items have been perceived plus (b) how many old ones have been recalled between its presentation and attempted recall. Thus, if an item appears in Position  $n$  from the end of a list and the subject attempts to recall it after recalling  $m$  other items, it is as if the item had appeared in position  $i = n + m$  in the list, and recall was attempted at the end of the list. This assumption is rather strong, but recent studies by

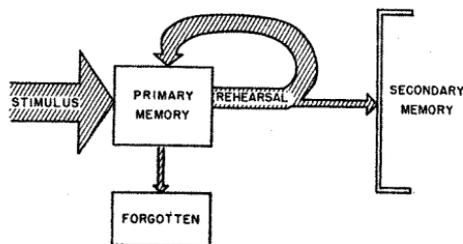


FIG. 2. The primary and secondary memory system. (All verbal items enter PM, where they are either rehearsed or forgotten. Rehearsed items may enter SM.)

Murdock (1963) and by Tulving and Arbuckle (1963) have, in fact, failed to reveal any consistent differences between stimulus- and response-induced interference in the retention of paired associates. It may not be unreasonable to suppose, therefore, that the two sources of interference exert equivalent effects on free and serial recall.

According to our hypothesis, then, the probability of recalling an item which has been followed by  $i$  subsequent items is given by the probability that it is in PM, in SM, or in both. Assuming that these probabilities combine independently,

$$R(i) = P(i) + S(i) - P(i)S(i) \quad [1]$$

where  $R(i)$  is the probability that Item  $i$  will be recalled,  $P(i)$  is the probability that it is in PM, and  $S(i)$  the probability that it is in SM. The probability that this item is in PM is then given by

$$P(i) = [R(i) - S(i)] / [1 - S(i)]. \quad [2]$$

We assume that  $P(i)$  is a monotonic decreasing function of  $i$  and that

$$\lim_{i \rightarrow \infty} P(i) = 0.$$

We should like specifically to test the hypothesis that  $P(i)$  is independent of the value of  $S(i)$  and, in fact, varies with  $i$  in the manner of the probe-digit data. (This hypothesis is stated more

formally in the Appendix.) In order to do so, we need data on verbal retention that meet the following requirements.

1. They should come from an experimental situation where at least some of the items are retrieved from PM.

2. The subject should have been allowed to rehearse, so that  $S(i) > 0$ .

3. The value of  $S(i)$  should preferably be constant and independent of  $i$ .

4. The experimental lists should be long enough to let us estimate  $S(i)$  for  $i > 12$ .

5. We should know the location of a given item in the stimulus list ( $n$ ) and in the recall list ( $m$ ), so as to be able to estimate the total number of interfering items ( $i = n + m$ ).

### Free Recall

The free-recall experiment is well suited to our purposes. Subjects can (and usually do) recall the last few items in a list right away, and the middle portion of the serial position curve (after the first three and before the last seven items) is effectively flat, thereby providing a convenient estimate of  $S(i)$  (Deese & Kaufman, 1957; Murdock, 1962; Waugh, 1962).

Testing our hypothesis against data collected in a free-recall experiment therefore involves the following steps:

1. First, we estimate  $S(i)$  from the average proportion of items recalled from the middle of a long list.

2. We then estimate  $P(i)$  for each of the last seven items in the list by Equation 2.

3. We plot this estimate against  $n + m = i$  and compare the resulting function with that shown in Figure 1.

Fortunately, we did not have to perform a free-recall experiment especially for this purpose: several such studies

have been carried out and reported in sufficient detail to enable us to test our hypothesis against their results. We have chosen to analyze four sets of data collected by three different investigators: Deese and Kaufman (1957), Murdock (1962), and two as yet unpublished experiments conducted by Waugh. The two principal variables that affect  $S(i)$  in free recall appear to be length of list (the amount of material that is to be retained) and presentation time (the amount of time available for the rehearsal of a given item). Manipulating these variables results in orderly changes in the value of  $S(i)$ , so that our estimates range from .08 to .45 across the four experiments.

1. In Deese and Kaufman's study, the subjects listened to lists of 32 unrelated English words read at a rate of one per second, and began recalling them immediately after the last had been spoken. Deese and Kaufman have presented a serial position curve based on these data and have also reported the relation between an item's serial position in recall and its position in the original list. We can thereby estimate  $i$  for each item in their lists, letting an item's average position in recall be our estimator of the amount of response interference ( $m$ ).<sup>3</sup> We estimated  $S(i)$  by the proportion of items recalled after the first three and before the last seven serial positions in the original list.<sup>4</sup> (This

<sup>3</sup> It is not really correct to use the average of the serial positions in recall as an estimate of  $m + 1$ : the total effect of response interference should depend on the variance of this distribution as well as on its mean or median. It is the only alternative open to us, however, since our correction for asymptote must be applied to the average proportion of items retained, estimated across serial position in recall.

<sup>4</sup> In estimating  $S(i)$ , we ignored the recall of the first three items on a list because

same general procedure will be followed in our subsequent analyses.)

The last seven points of Deese and Kaufman's serial position curve, taken from their Figure 1 and corrected for asymptote according to Equation 2, are plotted as a function of  $i$  in Figure 3. The dashed lines in Figure 3 represent the 99% confidence limits for the probe-digit function: a standard error for each point was estimated across subjects and experimental sessions. The uncorrected data are shown in Table 1.

2. Waugh's experiments were concerned with determining the number of items freely recalled from long lists as a function of presentation time. In her first experiment, 24, 30, 40, 60, or 120 different monosyllabic English words were read to the subjects at a rate of one per second. The proportion of items recalled varied inversely with list length, so that for each length of list there is a different serial position function. The asymptotes of these functions range from approximately .08 to .20. Median serial position in recall ( $m + 1$ ) was calculated for each of the last six items in a list;

they invariably show a primacy effect, perhaps the result of selective attention and rehearsal.

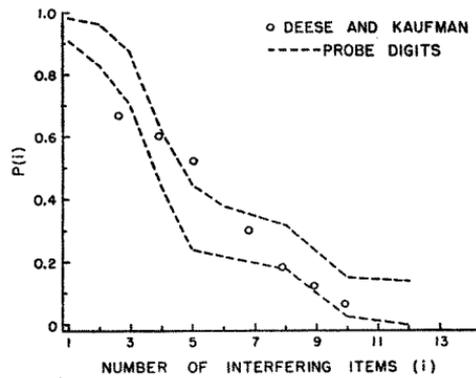


FIG. 3. Free-recall data from Deese and Kaufman (1957), corrected for asymptote and response interference.

Figure 4 shows  $S(i)$  as a function of  $i$  for each of these items. The uncorrected data appear in Table 2.

In Waugh's second experiment, the subjects listened to 30 different words presented at a rate of 1, 2, 3, 4, or 6 seconds per word. In each case the presentations were either massed—that is, each word was read one, two, or three times in a row, at a rate of one word per second or of one word every two seconds—or they were distributed—each word was read once at one, two, three, four, or six different places in a list, at a rate of one word per second. The results of this experiment indicate that whether the repetitions are massed

TABLE 1  
PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION  
OF SERIAL POSITION AND TOTAL TIME PER LIST

Number of interfering items	List length $\times$ seconds per item						
	32 $\times$ 1	40 $\times$ 1	20 $\times$ 2	30 $\times$ 1	15 $\times$ 2	20 $\times$ 1	10 $\times$ 2
0	.72	.96	.95	.97	.97	.96	.95
1	.67	.85	.88	.89	.88	.84	.83
2	.60	.71	.75	.74	.80	.76	.71
3	.42	.51	.57	.52	.62	.62	.67
4	.32	.40	.43	.39	.58	.39	.58
5	.27	.27	.38	.33	.49	.30	.45
6	.22	.22	.38	.24	.42	.26	.45
6+*	.17	.12	.27	.19	.38	.15	.45

Note.—Deese and Kaufman (1957), Column 1; Murdock (1961), Columns 2–6.  
\* Entries in this row represent the asymptotic value of  $R(n)$ .

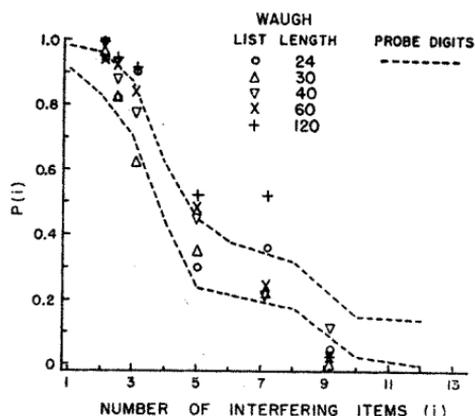


FIG. 4. Free-recall data from Waugh corrected for asymptote and response interference.

or distributed is of no importance; the probability that a word will be recalled is determined simply by the total number of seconds for which it is presented. Since this probability increases as a negatively accelerated function of presentation time, the asymptotic values of the serial position function obtained in this experiment ranged from approximately .14 (for 30 words each read once) to .45 (for 30 words each read six times). Average serial position in recall was again calculated for each of the last six items in a list. The retention functions for massed and distributed repetitions, cor-

TABLE 2

PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION OF STIMULUS INTERFERENCE AND NUMBER OF ITEMS PER LIST

Number of interfering items	List length				
	24	30	40	60	120
0	.95	.97	1.00	.95	1.00
1	.85	.85	.90	.93	.95
2	.92	.69	.81	.86	.92
3	.42	.46	.51	.53	.57
4	.47	.35	.31	.32	.57
5	.21	.17	.22	.14	.14
5+ <sup>a</sup>	.15	.17	.16	.12	.08

<sup>a</sup> Entries in this row represent the asymptotic value of  $R(n)$ .

rected for asymptote and response interference, are shown in Figures 5 and 6, respectively, along with the PM function obtained in our probe-digit experiment. The uncorrected data are shown in Table 2.

3. In Murdock's experiment, the subjects listened to lists of 20, 30, or 40 words read at a rate of 1 word per second and to lists of 10, 15, and 20 words read at a rate of 1 word every 2 seconds. Murdock found, as has Waugh (1963), that the probability of recalling a word that has been

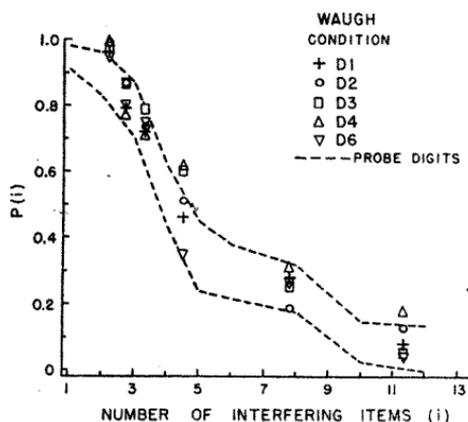


FIG. 5. Free-recall data from Waugh corrected for asymptote and response interference (1-6 distributed presentations per word).

listened to for 2 seconds is almost exactly twice the probability of recalling a word that has been listened to for 1 second. Murdock's data can therefore be grouped into three pairs of serial position curves: 10 words read at a rate of 1 every 2 seconds versus 20 words read at a rate of 1 per second; 15 words read at a rate of 1 every 2 seconds versus 30 read at a rate of 1 per second; and 20 words read at a rate of 1 every 2 seconds versus 40 read at a rate of 1 per second. Within each pair, there are two asymptotes, one of which is approximately twice the value of the other.

We have corrected Murdock's curves for asymptote—that is, for  $S(i)$ —and since he did not calculate serial position in recall for his words, we have plotted these corrected values of  $P(i)$  against the average values of  $i$  calculated by Waugh for words recalled under similar conditions in the experiment just described (see Figures 5 and 6).<sup>5</sup> Murdock's uncorrected data are shown in Table 1.

It is clear that an appreciable number of the points displayed in Figures 3 through 7 fall outside the confidence limits we have set for the probe-digit function. In general, the discrepancies between theoretical and observed values of  $P(i)$  appear to be unsystematic. They may have resulted from either of two possible sources which would not be reflected in the variance of the probe-digit function.

In the first place, we assume that  $S(i)$  is constant for all  $i$ . While  $S(i)$  does not in fact seem to vary systematically with  $i$  in the middle of a list, individual words do differ greatly in their susceptibility of storage in secondary memory: the serial position function for free recall is haphazardly jagged rather than perfectly flat. Thus, even one anomalously easy word in Location  $n$ , for instance, can greatly inflate our estimate of  $R(n)$  and hence  $P(n)$ . The probe-digit data would presumably not be subject to this kind of variability.

A second source of errors may lie in our estimation of  $i$ , or  $m+n$ . We have used average position in recall—call it  $\bar{m}+1$ —as our estimate of  $m+1$ . Even a small error in this estimate can lead to a sizable discrepancy between a theoretical and an observed value of  $P(i)$ , especially around the steep early portion of the function.

<sup>5</sup> The asymptotes for Murdock's curves were obtained by complementing his tabulated values for  $v$  (shown in his Table 2).

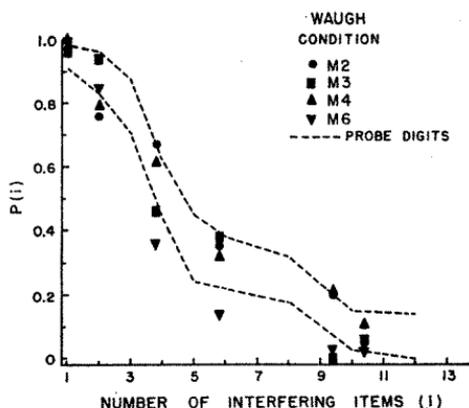


FIG. 6. Free-recall data from Waugh corrected for asymptote and response interference (2-6 massed presentations per word).

Errors of this sort would be reflected in Figures 4-7, where  $i$  and  $P(i)$  are derived from either partially or completely independent sets of data (in Figures 4-6 and Figure 7, respectively). Furthermore, we should in any case expect some discrepancy on purely mathematical grounds between  $P(i)$ , where  $(i)$  is the mean of a point distribution, as in the probe-digit experiment, and  $P(\bar{m}+n)$ , where  $m$  can assume any of a number of values, as in the free-recall data we have analyzed. Unfortunately, we are un-

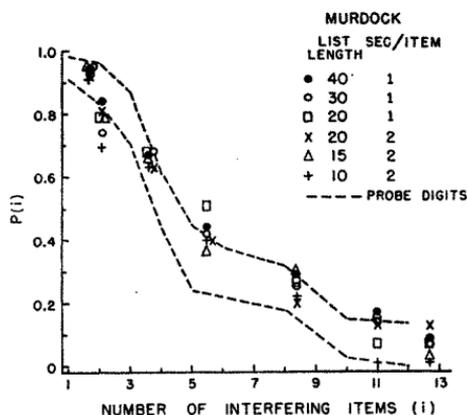


FIG. 7. Free-recall data from Murdock (1961), corrected for asymptote and response interference.

able to specify the magnitude of this expected discrepancy.

In view, therefore, of the likelihood of the errors we have just described, we believe that the fit between the probe-digit function and the free-recall data is fairly good and is, in fact, probably too close to attribute to chance. Actually, in one respect it is surprising that the probe-digit function should describe the free-recall data as well as it does. The probe-digit experiment tested the retention of digit pairs, whereas the free-recall experiments tested the retention of individual items. How are we justified in equating the two? One possibility is to assume that in the probe-digit experiment the subjects perceived and stored the digits as a series of overlapping pairs, rather than as single digits. In this case, the measure of interference would be given by the number of digit-pairs that follow any given pair, which is, of course, equal to the number of single digits that follow it. In the free-recall experiment, on the other hand, the subjects may have perceived the words as independent units, and the effective interference would then consist of single words, as we have in fact been assuming. The problem, then, can be restated as follows: why do pairs of digits and single words exert equal amounts of retroactive interference on like items in primary memory? There is little in the existing literature that sheds much light on this point.

### *Paired Associates*

Our model should, of course, be able to describe ordered as well as free recall. We face serious problems, however, in attempting to apply it to serial learning: if a list is long enough to furnish a stable estimate of  $S(i)$ , the probability that a given item will be in PM at the time of testing is negli-

gible, since serial items are customarily tested in the order in which they were presented. We must therefore turn to paired associates. In a recent study, Tulving and Arbuckle (1963) systematically varied the positions of the items on the recall list, and we have therefore applied our hypothesis to their data in the manner described above.

Tulving and Arbuckle presented number-word pairs to their subjects and then tested for the recall of each word by presenting only the number with which it had been paired. They were interested in measuring probability of recall after one trial as a function of an item's serial position in both the original list and the test list. We have estimated  $S(i)$  by averaging the recall probabilities for  $i > 13$ , excluding Items 1 and 2. The value of their serial position curve is fortunately constant in this region, as it was for free recall. Note that in this task, each pair presented after a given number and before the cue for its recall actually consists of *two* interfering items: a word plus a number. We have counted all items occurring between the test item and its recall—including the test number—as interfering items. We have analyzed the proportion of items presented in Positions 1 through 6 from the end of the stimulus list and tested in Positions 1 through 6 of the response list. These proportions are shown in Tulving and Arbuckle's Tables 2 and 4; we have pooled those that correspond to a given value of  $i$ . Thus  $i$ , or  $n + m$  (where  $n = j$  and  $m = i - j$ ), ranges from 1 to 11. These data are presented in Figure 8, along with our own estimate of  $P(i)$ . Again, considering the variability of  $S(i)$  that is not taken into account by our model, the fit between data and theory appears to be reasonably good.

In sum, then, we believe we can say

that the similarity between our probe-digit function and the various other, initially disparate, serial position curves shown in Figures 3-8 is consistent with the hypothesis that there is a primary memory store that is independent of any longer term store. The capacity of the primary store appears to be invariant under a wide variety of experimental conditions which do, however, affect the properties of the longer term store.

### Single-Item Retention

Much of the experimental work on memory in the past 5 years has focused on measuring the retention of a single verbal item—or of a brief list of items—over short intervals. A widely used procedure which was introduced by Peterson and Peterson (1959) is to expose an item (for example, a meaningless three-letter sequence) to a subject; have him perform some task that presumably monopolizes his attention (such as counting backwards by threes) for a specified number of seconds; and, finally, at the end of this interval, have him attempt to recall the critical item. The universal finding has been that retention decreases monotonically with the length of the retention interval. It has generally been assumed that the subject does not rehearse during the retention interval, that a number spoken by him does not interfere with a trigram previously spoken by the experimenter, and that therefore the observed decline over time in the retention of such an item reflects the pure decay of its memory trace. This general conclusion is clearly inconsistent with our results, since we have found that the length of the retention interval as such—within the limits we tested, naturally—is of relatively little importance in determining retention loss.

In seeking for a way to account for

this discrepancy, it occurred to us to question the assumption that, in an experiment of the sort described above, the numbers spoken by the subject during the retention interval do not interfere with the memory trace of the item he is supposed to retain. Some experimenters have, after all, reported that dissimilar items seem to interfere with one another just as much as do similar ones in the immediate recall of very short lists (Brown, 1958; Pillsbury & Sylvester, 1940). What would happen, therefore, if we were to define a three-digit number uttered by a subject in the course of a simple arithmetic calculation—counting backwards—as one unit of mnemonic interference? Could our model then describe the forgetting of single items over brief intervals? We have attempted to fit the data of two experimenters, Loess (in press) and Murdock (1961), by converting the retention interval into a corresponding number of interfering items. Murdock's subjects were trained to count at a steady rate of one number per second, so the number of interfering items in his experiment is equal to the retention interval in seconds. Loess' subjects counted at a rate of one number every 1.5 seconds; we have therefore multiplied the length of his

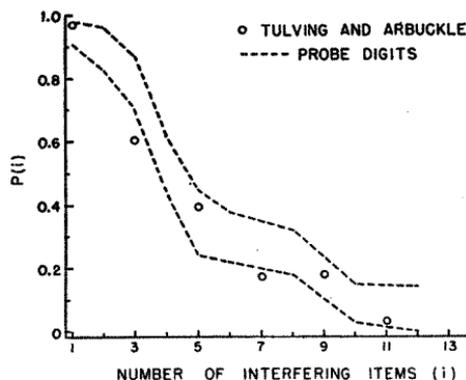


FIG. 8. Paired-associate data from Tulving and Arbuckle (1963) corrected for asymptote and response interference.

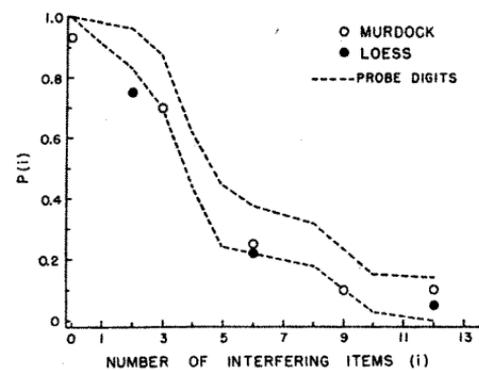


FIG. 9. The retention of three-item lists compared with the probe-digit function, (Loess' data denote the proportion of consonant trigrams recalled after various retention intervals; Murdock's data represent the average proportion of trigrams and word triads retained after a given interval.)

retention intervals by  $\frac{2}{3}$  in order to obtain the equivalent number of interfering items. We estimate  $S(i)$  in both cases by the relative frequency of recall at  $i = 18$ .<sup>6</sup>

The two sets of data, corrected for asymptote, are shown in Figure 9, along with the probe-digit function. The correspondence between them is reasonably close. It is possible, of course, that this agreement between theory and fact is simply a matter of luck, depending, as it does, on the arbitrary assumption that a three-digit number generated by the subject himself is psychologically equivalent to a one-digit number presented by the experimenter during the retention interval (as in the probe-digit study). Obviously we cannot draw any firm conclusions about the effect of interference on the retention of single items until this assumption is justified empirically. We can only point out that

<sup>6</sup> We have also tried to analyze the results of Peterson and Peterson (1959) but without success. Part of the difficulty may result from the fact that their subjects may not have adhered strictly to a prescribed rate of counting during the retention interval (L. Peterson, personal communication, 1964).

the results of Murdock and Loess do not necessarily contradict our model.

## DISCUSSION

We should at this point like to consider the general question of whether all verbal information is stored in the same system or whether, as we have assumed here, there are two independent mnemonic processes that contribute to retention even over very short intervals. The proponents of a unitary theory of memory, eloquently led by Melton (1963), have argued that recall after a few seconds is affected in very similar ways by the variables that govern recall over much longer intervals; and that therefore the distinction between a short-term memory mechanism, on the one hand, and a longer term mechanism, on the other, is purely arbitrary. The following facts have been cited in support of this argument:

1. Short-term retention improves, just as does long-term retention, when the material to be recalled is repeated before a test of retention, or when it is repeated between successive tests (Hebb, 1961; Hellyer, 1962).
2. Retention after a brief delay is subject to proactive interference, as is retention after a long delay (Keppel & Underwood, 1962; Loess, in press). Why, asks the unitary theorist, should we distinguish between short- and long-term retention if we cannot find any quantitative and experimentally manipulatable differences between them? This question might well be disturbing if one took the position that the two processes have sharply defined non-overlapping temporal boundaries such that items recalled within some critical interval after their initial occurrence must have been retrieved from one system, whereas items recalled beyond this interval must have been retrieved

from another. (Such a view would imply, interestingly enough, that an item would have to remain in a short-term storage for some specified number of seconds before passing into longer term storage, if it did so at all.)

But what if we do not require that the two systems be mutually exclusive? Then the probability that an item will be recalled will depend on both the probability that it is still in PM and the probability that it has entered into SM in the interval between its presentation and the start of the interfering sequence (or even during this sequence, if the subject is able to rehearse). All those variables that determine  $S(i)$  for a given item—such as its position in a closely spaced series of tests, or the number of times it has been repeated—will then determine the observed proportion recalled after a brief interval. We believe we have shown, however, that  $P(i)$  depends only on  $i$  and remains invariant with changes in  $S(i)$ ; and we submit that most of the published data on short-term retention actually reflect the properties of both memory systems.

We would like to make one final point: the existence of some rather compelling introspective evidence in favor of two distinct mnemonic sys-

tems. PM, as we have defined it here, is best illustrated by a person's ability to recall verbatim the most recent few words in a sentence that he is hearing or speaking, even when he is barely paying attention to what is being said, or to what he is saying. Given that the flow of speech is intelligible, failures in the immediate recall of words we have just heard—errors of either omission, transposition, or substitution—are probably so rare as to be abnormal. Indeed, we believe that it would be impossible to understand or to generate a grammatical utterance if we lacked this rather remarkable mnemonic capacity. In order to recall a sentence verbatim at a later time, however, we usually have to rehearse it while it is still available in PM.

The same effect holds for meaningless arrangements of verbal items. If we present a subject with a random string of words, letters, or digits, and ask him to reproduce them in any order he chooses, he can maximize the number he recalls by "unloading" the last few items immediately. Most subjects in free-recall experiments report that these very late items tend to be lost if they are not recalled immediately, whereas items that came earlier in the list can be retrieved at leisure, if they

TABLE 3  
PROPORTION OF ITEMS FREELY RECALLED AS A FUNCTION OF  
STIMULUS INTERFERENCE AND PRESENTATION TIME

Number of intervening items	Seconds per item									
	Distributed					Massed				
	1	2	3	4	6	2	3	4	6	
0	.96	.99	.97	1.00	.97	.98	.97	1.00	1.00	
1	.82	.90	.91	.86	.89	.82	.96	.87	.91	
2	.76	.81	.86	.82	.87	.75	.63	.76	.63	
3	.54	.64	.73	.76	.65	.51	.58	.58	.50	
4	.38	.40	.50	.57	.60	.40	.31	.51	.44	
5	.21	.36	.36	.49	.48	.27	.36	.45	.44	
5+ <sup>a</sup>	.14	.26	.32	.38	.45	.25	.31	.38	.42	

<sup>a</sup> Entries in this row represent the asymptotic value of  $R(n)$ .

can be recalled at all. In the colorful terminology of one such subject (Waugh, 1961), the most recent items in a verbal series reside temporarily in a kind of "echo box," from which they can be effortlessly parroted back. When an experienced subject is trying to memorize a list of serial items, moreover, he "fills up" successive echo boxes as the list is read to him and attempts to rehearse the contents of each. He will invariably lose some items if rehearsal is delayed too long or if he attempts to load his echo box with more items than it can hold. We think it very likely that the PM function describes the (variable) capacity of this mechanism. We would remind you in this connection that, within very broad limits, the rate at which someone is speaking does not affect your ability to follow his words—just as differences in the rate at which meaningless lists of digits are presented do not exert any profound effect on the PM function.

### CONCLUSIONS

We have tried to demonstrate the existence of a short-term or PM system that is independent of any longer term or secondary store by showing that one function relating probability of recall to number of intervening items can describe a number of seemingly disparate sets of experimental results. In doing so, we have deliberately avoided discussing a number of problems raised in our analyses. Foremost in our list of problems is the definition of an item.

Certainly the idea of a discrete verbal unit is crucial to our theory. The interference effect that we have studied seems to be invariant over a broad class of units and combinations of units—single digits, nonsense trigrams, and meaningful words. How long a string of such primitive units can we combine and still have one item? Is an item determined by our grammatical habits? Is it determined by the duration of the verbal stimulus? Is it determined by both? We do not know.

We have also avoided discussing the possible rules whereby items now in PM are displaced by later items. Are items lost independently of one another, or do they hang and fall together? It may perhaps prove difficult to answer this question experimentally, but it should not be impossible.

Finally, at what stage in the processing of incoming information does our PM reside? Is it in the peripheral sensory mechanism? Probably not. The work of Sperling (1960) indicates that "sensory memory"—to use Peterson's (1963) phrase—decays within a matter of milliseconds, whereas we have dealt in our analysis with retention intervals on the order of seconds. Does storage in PM precede the attachment of meaning to discrete verbal stimuli? Must a verbal stimulus be transformed into an auditory image in order to be stored in PM, even if it was presented visually? We refer the reader to a recent paper by Sperling (1963) for some thoughts on the latter question.

### APPENDIX

A formal discussion of the interaction between PM and SM can be provided by a simple three-state Markov process. The assumptions of the model are:

1. There are three states of memory:  $S$ ,  $P$ , and the null state,  $G$ .

2. The probability of recalling an item from either State  $S$  or State  $P$  is unity: items cannot be recalled from the null state, but they may be guessed with Probability  $g$ .

3. Items can only pass into State  $S$  when they are rehearsed and, for the ex-

periments discussed in this paper, we assume that items are rehearsed only when they are presented. The probability that an item is stored in  $S$ , given that it was successfully rehearsed, is  $\alpha$ .

4. Items in  $P$  are interfered with by later presentation of different items: the probability that an item returns to the null state on the presentation of the  $i$ th interfering item is  $\delta_i$ .

The following equivalents hold between the terms defined for the Markov model and the terms defined in the body of the paper:

1.  $P_i(S)$  is equivalent to  $S(i)$ .
2.  $P_i(P)$  is equivalent to  $P(i)[1 - S(i)]$ .
3.  $\delta_i$  is equivalent to  $1 - P(i)$ .

Now, define the random variable  $\pi$  with Value 1 if the test item is presented, and with Value 0 if some other (interfering) item is presented. (We can also let  $\pi$  be a probability—namely, the probability that the test item is presented. The formal statement of the model does not change with this redefinition.)

The transition probabilities for any given stimulus item (the test item) are specified by the matrix

$$\begin{array}{c} S \qquad P \qquad G \\ \begin{array}{l} S \\ P \\ G \end{array} \begin{bmatrix} 1 & 0 & 0 \\ \alpha\pi & (1-\pi)(1-\delta_i) + \pi(1-\alpha) & (1-\pi)\delta_i \\ \alpha\pi & (1-\alpha)\pi & 1-\pi \end{bmatrix} \end{array}$$

Unfortunately, it is difficult to work with transition matrices of this form (with time-varying parameters). One approximation would be to let  $\delta_i = \delta$ , independent of  $i$ . This approximation yields an exponential decay function of the form  $P_i(P) = (1-\alpha)(1-\delta)^{i-1}$ . This is clearly not correct for the results of our experiment (Figure 1); but, for some purposes, it may not be a bad approximation. A model very similar mathematically to that produced by this simple approximation for  $\delta_i$  has been studied by Atkinson and Crothers (1964), who found it to be quite good for certain types of paired-associates experiments. Their model, however, is derived from quite different considerations.

For any experiments with controlled rehearsals, the probability that an item reaches State  $S$  (or SM) is completely independent of the properties of the short-term state ( $P$  or PM). This is true because, as far as State  $S$  is concerned, the general transition matrix can be reduced by combining States  $P$  and  $G$  to form the "lumped" State  $P'$ . The new matrix is

$$\begin{array}{c} S \qquad P' \\ \begin{array}{l} S \\ P' \end{array} \begin{bmatrix} 1 & 0 \\ \alpha\pi & 1-\alpha\pi \end{bmatrix} \end{array}$$

This is a simple one-element Markov model. This means that although the complete description of the verbal learning process requires a description of the short-term state, a study of only the long-term retention of items can ignore the short-term memory.

## REFERENCES

- ATKINSON, R. C., & CROTHERS, E. J. A comparison of paired-associate learning models having different acquisition and retention axioms. *Journal of Mathematical Psychology*, 1964, 1, 285-315.
- BROADBENT, D. E. *Perception and communication*. New York: Pergamon Press, 1958.
- BROWN, J. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 1958, 10, 12-21.
- BUGELSKI, B. R. Presentation time, total time, and mediation in paired-associate learning. *Journal of Experimental Psychology*, 1962, 63, 409-412.
- CLARK, L. L., LANSFORD, T. E., & DALLENBACH, K. M. Repetition and associative learning. *American Journal of Psychology*, 1960, 73, 22-40.
- CONRAD, R. Decay theory of immediate memory. *Nature*, 1957, 179, 831-832.
- DEESE, J., & KAUFMAN, R. A. Sequential effects in recall of unorganized and sequentially organized material. *Journal of Experimental Psychology*, 1957, 54, 180-187.
- HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
- HEBB, D. O. Distinctive features of learning in the higher animal. In J. F. Delacresnay (Ed.), *Brain mechanisms and learning*. London: Oxford Univer. Press, 1961. Pp. 37-46.

- HELLYER, S. Supplementary report: Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology*, 1962, **64**, 650.
- JAMES, W. *The principles of psychology*. Vol. 1. New York: Holt, 1890. Ch. 16.
- KEPPEL, G., & UNDERWOOD, B. J. Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 153-161.
- LOESS, H. Proactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, in press.
- MCGEOCH, J. A. Forgetting and the law of disuse. *Psychological Review*, 1932, **39**, 352-370.
- MELTON, A. W. Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 1-21.
- MURDOCK, B. B., JR. The retention of individual items. *Journal of Experimental Psychology*, 1961, **62**, 618-625.
- MURDOCK, B. B., JR. The serial position effect in free recall. *Journal of Experimental Psychology*, 1962, **64**, 482-488.
- MURDOCK, B. B., JR. Interpolated recall in short-term memory. *Journal of Experimental Psychology*, 1963, **66**, 525-532.
- PETERSON, L. R. Immediate memory: Data and theory. In C. N. Cofer, & Barbara Musgrave (Eds.), *Verbal behavior and learning: Problems and processes*. New York: McGraw-Hill, 1963. Pp. 336-353.
- PETERSON, L. R., & PETERSON, M. J. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 1959, **58**, 193-198.
- PILLSBURY, W. B., & SYLVESTER, A. Retroactive and proactive inhibition in immediate memory. *Journal of Experimental Psychology*, 1940, **27**, 532-545.
- POSNER, M. I. Immediate memory in sequential tasks. *Psychological Bulletin*, 1963, **60**, 333-349.
- POSTMAN, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 152-179.
- SPELRING, G. The information available in brief visual presentations. *Psychological Monographs*, 1960, **74** (11, Whole No. 498).
- SPELRING, G. A model for visual memory tasks. *Human Factors*, 1963, **5**, 19-36.
- TULVING, E., & ARBUCKLE, T. Y. Sources of intratrial interference in immediate recall of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 1963, **1**, 321-334.
- UNDERWOOD, B. J., & KEPPEL, G. An evaluation of two problems of method in the study of retention. *American Journal of Psychology*, 1962, **75**, 1-17.
- WAUGH, N. C. Free versus serial recall. *Journal of Experimental Psychology*, 1961, **62**, 496-502.
- WAUGH, N. C. The effect of intralist repetition on free recall. *Journal of Verbal Learning and Verbal Behavior*, 1962, **1**, 95-99.
- WAUGH, N. C. Immediate memory as a function of repetition. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 107-112.

(Received March 10, 1964)