JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 17, 309–324 (1978)

Controlled Rehearsal in Single-Trial Free Recall[1]

BENNET MURDOCK AND JANET METCALFE

University of Toronto, Toronto, Ontario M5S 1A1, Canada

Interpretation of data from the overt-rehearsal procedure may be beclouded by item-selection artifacts. To test this hypothesis, a controlled rehearsal procedure was used where to-be-rehearsed items were presented to the subject rather than selected by him. Which items were rehearsed and how many were comparable under the controlled- and the overt-rehearsal conditions. Total recall and the nominal and functional serial-position curves were similar under the two conditions, suggesting that item-selection artifacts could not be too serious. During Session 5 earlier lists were used that the subject had seen under the overt-rehearsal condition, and, using the controlled-rehearsal method, the original rehearsal order was preserved or scrambled. No differences in performance were found. During Session 6 covert rehearsal was compared to the overt rehearsal previously used, and again no differences were found. During Session 7, the rehearsal pattern with the controlled-rehearsal procedure was manipulated and it was found that recall deteriorated when a poor rehearsal pattern was used. These results suggest that a controlled-rehearsal procedure may be useful to study rehearsal processes in free recall, and a strength–decay model was outlined to try to explain the origin of nominal and functional serial–position curves.

One of the relatively solid principles of single-trial free recall has been that rate of presentation affects the prerecency but not the recency component of the serial-position curve. Recently Brodie and Murdock (1977) questioned this conclusion. They showed that there was a large difference in the recency part of the curve when functional rather than nominal serial-position curves were plotted. A functional serial-position curve plots recall probability as a function of where an item was last rehearsed rather than where it was last presented. This finding has rather important implications for our understanding of single-trial free recall, since it seems to contradict encoding models (e.g., Craik & Lockhart, 1972), storage models (e.g., Atkinson & Shiffrin, 1968; Glanzer, 1972), and retrieval models (e.g., Tulving, 1968). As detailed in Brodie and Murdock, all these models would seem to predict the classical presentation-rate effect for the functional serial-position curves, yet the opposite effect occurred.

Functional serial-position curves were obtained by using the overt-rehearsal procedure (Corballis, 1969; Rundus & Atkinson, 1970). The overt-rehearsal procedure, however, has interpretive problems, the main one being item selection effects. Some items may be easier to remember than other items; these items are rehearsed more. So, it is not that rehearsal "strengthens" items and makes them easier to remember; rather, those items which are easier to remember may be rehearsed more. The relationship is correlational, not causal. If this problem is serious enough to invalidate the use of the overt rehearsal procedure, then functional serial-position curves are likewise suspect and the Brodie and Murdock conclusion is questionable.

More specifically, carry this argument to its logical extreme. Suppose the easiest items are rehearsed the most and the hardest items the least. Then, on average, the position of the last

[1] This research was supported by Grant APA 146 from the National Research Council of Canada and OMHF 164 from the Ontario Mental Health Foundation. The authors would like to thank Howard Kaplan and Peter Liepa for programming assistance and Del Brodie for many helpful comments.

0022–5371/78/0173–0309$02.00/0

rehearsal of any given item will tend to reflect how easy it is to remember (the easier, the later). Then, what the functional serial-position curve is showing us is a curve of item difficulty. The easiest items are best recalled and the hardest items least recalled, hardly a surprising result. Furthermore, by this reasoning one can even explain the presentation-rate effect. The slower the presentation time, the more chance the subject has to sort out the easier from the harder items, so the sharper the gradient; or (or in addition) the slower presentation rate, the greater the total number of rehearsals, so the rehearsal gradient is steeper as the presentation rate becomes slower. Since item difficulty projects onto both the last rehearsal of a given item and the total number of rehearsals of a given item, the differences in the functional serial-position curves between fast and slow presentation rates are then attributable to nothing more than item-selection artifacts.

Brodie and Murdock were not unaware of item-selection effects; in fact, their Experiment II was addressed particularly to this issue. Subjects were called back 7 weeks or more after the conclusion of Experiment I to be retested. Experiment I lists were repeated either in the same order as they had been shown originally or in the order that the subject had last rehearsed each item. The fact that there were no differences in the recall of these two kinds of lists was taken as an indication that item-selection artifacts in the overt-rehearsal procedure could not be too serious. This is, however, a weak test of the issue since one is really only manipulating presentation order and not rehearsal order.

A stronger test would be to have a procedure which is completely free of item-selection artifacts. That is what we attempted to do here, and the method involved the use of a "controlled rehearsal" procedure. Rehearsed items were randomly selected according to certain constraints and not self-selected by the subject. During each rehearsal interval (i.e., the time between presentation of Item i and Item $i + 1$), rehearsal items were presented to the subject. These rehearsal items were presented one at a time, and were randomly selected from Item 1, Item 2, ..., Item i. The subject was instructed "to think about" only the item currently in view.

If this controlled-rehearsal procedure is to be a realistic approximation to the overt-rehearsal procedure, then the randomization routine must generate a rehearsal pattern characteristic of the subject being tested. To this end, the randomization was tailor-made for each individual subject. We matched on two dimensions: the number of items rehearsed and their serial positions. Obviously, it was necessary to have some data as a baseline, so each subject ran one or two overt-rehearsal sessions prior to the controlled-rehearsal sessions. Further details on the method will be described under the Procedure section.

The main purpose, then, of this controlled-rehearsal procedure was to assess the importance of item-selection artifacts in the overt-rehearsal procedure. A further but different test of the same issue was provided by the Session 5 data. There were two sessions of overt rehearsal and two sessions of controlled rehearsal. Session 5 followed these four. We waited 7 weeks or more after Session 4 and then retested the subjects on the same lists they had seen on Session 1 (overt rehearsal). Session 5 was run under the controlled-rehearsal conditions; half the lists were presented and "rehearsed" in the Session 1 order; the other half were rehearsed in the same order but the presentation order was scrambled. Thus, the rehearsal pattern was exactly what the subject had used on that list, but the item order was different. We know (Brodie & Murdock, 1977, Expt II) that, with retesting, subjects show no savings for repeated lists. Thus, a comparison of these two conditions (same order versus scrambled order) will provide additional information on the importance of item-selection artifacts.

The two controlled-rehearsal sessions included a final free recall, where subjects were instructed to recall as many items as possible from the entire session. The final free recall was included to obtain a measure of performance over a longer retention interval than the immediate recall. Also, several additional sessions (following Session 5) were run to test some of conclusions suggested by the preplanned experiment. These will be reported after the main results have been presented.

METHOD

Subjects

The subjects were 16 undergraduate students of both sexes and were recruited by an advertisement. They were paid for their services.

Materials

The materials comprising each list were random samples from the Toronto word pool, a list of 1040 two-syllable common English words not more than eight letters long with homophones, contractions, archaic words, and proper nouns deleted. Lists were always 20 items long, and there were 16 lists per session. To facilitate scoring, the same 16 lists were used for all subjects on each of the overt-rehearsal sessions. No words were repeated in the 32 total (2 × 16) lists. Order of presentation of lists within each session was counterbalanced across subjects in a 16 × 16 Latin Square. For the two controlled-rehearsal sessions, each list was a random sample without replacement from the word pool.

Procedure

Basically, there were two different presentation conditions, and we shall refer to them as the overt-rehearsal condition and the controlled-rehearsal condition. The overt-rehearsal condition was modeled directly after Rundus and Atkinson (1970). We used a slow presentation rate (one word every 5 seconds) with visual presentation. Each word was in view for 1 second, and the screen was blank for 4 seconds. The instructions followed Brodie and Murdock (1977); namely, the subject was instructed to report continually which word he was thinking about, and no restrictions were placed on which words he should think about or how he should try to encode them. Each subject was run individually, and the rehearsal data were tape-recorded for subsequent analysis. Recall was written, and the recall period for each list was self-paced.

The controlled-rehearsal condition was run on a PDP-12A laboratory computer with the words displayed on the CRT. Each "list" word was presented near the top of the CRT display and was on for 1 second. When the word went off, the rehearsal words were presented for the next 4 seconds near the bottom of the CRT display. Thus, instead of the self-selection inherent in the overt-rehearsal procedure, the randomization routine of the computer determined what the subject rehearsed. Naturally, the subject was fully informed about all this and told to think only about each word currently in view. The fact that there would be 20 list words and some variable number of rehearsal words was included in the explanation. As a control for any possible modality effect, the subject was instructed to say all words aloud.

With this procedure we wanted to mimic the rehearsal pattern of each subject as closely as possible. The two variables we matched on were (a) the distribution of the number of words rehearsed in each 4-second interval; and (b) the lag distribution of the rehearsed words. For each subject, we took his previous overt-rehearsal session and computed the distribution of number of words rehearsed (summed over lists and serial position). This distribution was then available for (a). For (b), we computed the lag distribution separately for each serial position (i.e., separately for Serial Position 1, 2, 3, ..., 20) but pooled over the 16 overt-rehearsal lists. A lag of 0 indicated rehearsal of the current list word, a lag of 1 the rehearsal of the list word one back, etc. The randomization routine in the computer then made use of these two distributions (i.e., a and b) to determine what to present to the subject as rehearsals.

More specifically, the randomization routine worked as follows. First, it determined x, the number of words to be rehearsed in the oncoming 4-second interval. Distribution (a) governed the probabilities associated with each outcome. Then, the oncoming rehearsal interval was divided into $4/x$ seconds so that each rehearsal word would have equal presentation time. The rehearsal words were determined by x successive draws (with replacement) from Distribution (b). As has been mentioned, whereas there was one Distribution (a) for all serial positions, there was a unique Distribution (b) for each of the 20 serial positions.

The Session 5 data were collected 7 weeks or more after Session 4 had been run. The purpose of this delay was to ensure that the subjects had forgotten (i.e., would show no savings from relearning on) the Session 1 lists. (Relevant evidence on this point may be found in Brodie & Murdock, 1977.) Session 5 was run under controlled-rehearsal conditions, and each list consisted of the same words seen by that subject on that list in Session 1. The randomization routine was replaced by the actual rehearsal data of that subject on that list in Session 1. Then, of the 16 lists in the session, half were scrambled and half were not prior to presentation. (Which lists were scrambled was randomly determined for each subject.) In this way, we can preserve the general rehearsal pattern for the subject but determine the rehearsal items. For the scrambled lists, the pattern is appropriate but the detailed item-rehearsal pattern is broken up. For the preserved lists, not only is the pattern the same but also the words themselves are the same.

Final free recall followed the two controlled-rehearsal sessions. Subjects were not given advance notification before the first session but did know before the second session. After the subject left the computer room, there was a short conversation with the experimenter; then the subject was asked to do the final free recall. The instructions were simply to write down as many words as possible from the entire experimental session. There was no final free recall after the two overt-rehearsal sessions.

Design

For control purposes, it would be desirable to counterbalance the order of conditions across subjects. Since, however, data from the overt-rehearsal condition were needed before the controlled-rehearsal sessions could be run, complete counterbalancing was impossible. A partial counterbalancing, however, was used. For half the subjects, Sessions 1 and 2 were overt rehearsal while Sessions 3 and 4 were controlled rehearsal. For the other half of the subjects, Sessions 1 and 4 were overt rehearsal while Sessions 2 and 3 were controlled rehearsal. Comparisons between these two counterbalancing groups will tell us whether there are practice effects to complicate the interpretation of the results.

RESULTS

The rehearsal frequencies for each presentation position in the overt-rehearsal sessions are shown in Table 1. These data are pooled over the 16 subjects and the 32 lists per subject. There was a strong recency effect at all presentation positions. There were also primacy effects, but they were not as marked. These data were used to determine Distribution b (the lag distribution) for the controlled-rehearsal sessions. [The distributions, of course, were not pooled over subjects; each subject had his own rehearsal distribution. Also, it was only based on the first session (Lists 1–16) and not both sessions.]

The most global measure of performance is the total number (or proportion) of words recalled under the two conditions. The proportions, summed over sessions, subjects, and lists were .495 for the overt-rehearsal condition and .457 for the controlled-rehearsal condition. Thus, on the average, subjects

TABLE 1

REHEARSAL FREQUENCIES FOR EACH PRESENTATION POSITION IN THE OVERT REHEARSAL SESSIONS

Presentation position																				
1	1564																			
2	878	1065																		
3	649	705	873																	
4	443	464	500	762																
5	335	322	339	424	700															
6	250	275	212	247	302	668														
7	193	192	179	167	185	270	657													
8	153	163	147	142	146	194	259	663												
9	134	143	112	117	117	116	130	283	665											
10	147	130	110	114	95	94	99	140	221	664										
11	137	133	104	73	77	65	57	111	139	194	663									
12	114	125	70	92	72	61	53	79	85	82	201	617								
13	129	110	105	72	71	65	42	67	60	65	105	224	681							
14	119	110	86	59	41	42	37	39	38	45	72	113	223	673						
15	105	83	70	65	50	39	29	56	45	34	58	92	129	237	667					
16	128	105	75	50	68	33	43	43	46	19	42	40	67	126	216	646				
17	110	101	71	66	46	31	30	31	29	29	33	48	48	90	120	202	684			
18	91	88	63	51	53	57	25	30	30	20	21	35	41	49	76	106	224	633		
19	93	89	72	42	44	30	26	35	33	23	32	26	32	41	55	61	122	216	693	
20	101	88	78	59	46	39	20	25	26	14	13	24	16	37	36	35	66	128	217	597

recalled 9.90 words under the overt-rehearsal condition and 9.14 words under the controlled-rehearsal condition. Even though this difference was small, an analysis of variance showed it to be highly significant; $F(1, \infty) = 41.1$, $p < .001$. All analyses of variance in this paper were done by the method suggested in Murdock & Ogilvie (1968). Under the assumption of binomial variability, the E(MS) value is $1/n$, where n (here 8) is the number of observations entering into each macroblock. Since the E(MS) value was used as the error term, the denominator of the F ratio is a theoretical and not a calculated value. This analysis pooled over sessions regardless of order; the 32 overt-rehearsal lists were compared with the 32 controlled-rehearsal lists.

To look at practice effects, since there is an overt-rehearsal versus controlled-rehearsal difference, it is necessary to make the comparison separately. Averaged over the two counterbalancing groups, for the overt-rehearsal condition the mean recall proportions were .482 and .508 for the first and the second session, respectively. For the controlled-rehearsal condition the mean recall proportions were .457 and .456 for the two sessions. Given these values, it would seem safe to conclude that practice effects do not pose an interpretive problem for the results of this experiment.

The nominal and functional serial-position curves for the overt-rehearsal condition are shown in Fig. 1, and the nominal and functional serial-position curves for the controlled-rehearsal condition are shown in Fig. 2. On casual inspection, at least, the similarities are quite striking. The general pattern of results in the two figures is overall much the same. Clearly the nominal and functional curves are different, and they replicate quite nicely previous results (e.g., Fig. 1 of Brodie & Murdock, 1977). Perhaps we should reiterate the difference between nominal and functional serial-position curves. In comparison to the nominal, the functional serial-position curves: (a) have no primacy effect, (b) have a questionable asymptote, and (c) have a more extended recency effect. They are easier to characterize analytically.

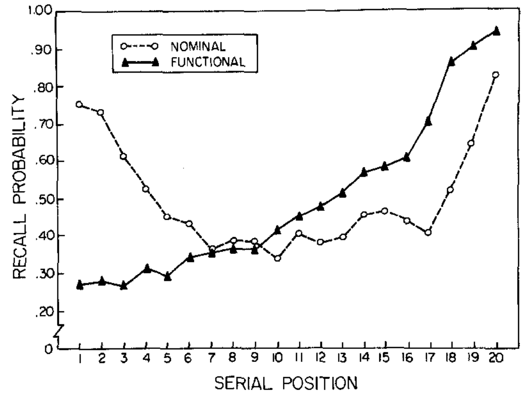

FIG. 1. Nominal and functional serial-position curves for the overt-rehearsal sessions.

One way to describe the functional serial-position curves is by means of the Lockhart attribute model, described in Murdock (1974, Pp. 43–45). According to this model, attributes of items are encoded and the parameter ϕ characterizes their number. Encoded attributes undergo fluctuation; α is the decay parameter while β is the recovery parameter. Every item which has at least one remembered attribute at the end of list presentation is recalled. Equation 2.17 (Murdock, 1974, P. 44) was used to obtain the predicted values, and the value of t for each item was obtained by subtracting its functional serial position

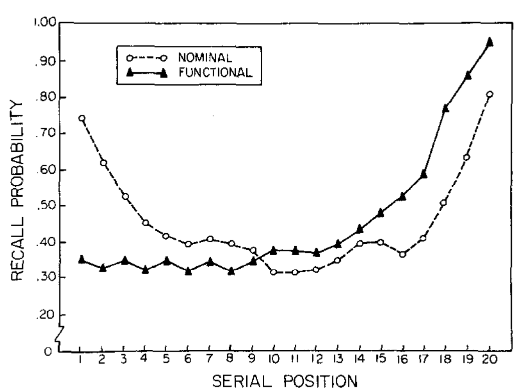

FIG. 2. Nominal and functional serial-position curves for the controlled-rehearsal sessions.

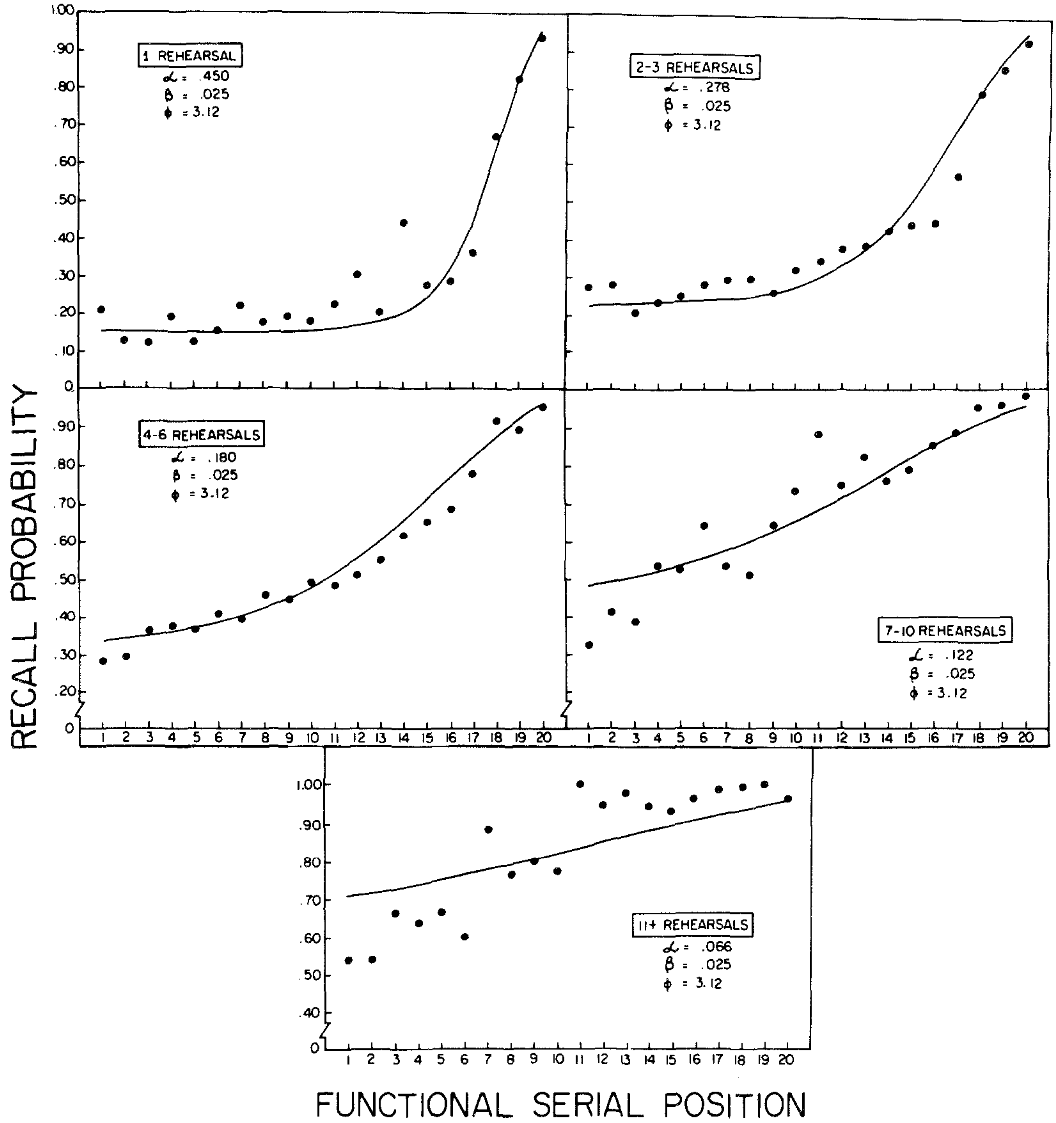

Fig. 3. Functional serial-position curves for the overt-rehearsal sessions partitioned on the number of rehearsals.

from 20. Thus, t was the backward functional serial position, with the last item having a value of zero. The data were fit to the model by SIMPLEX (Nelder & Mead, 1965) and the results are shown in Figs. 3 and 4. Figure 3 shows the results for the overt-rehearsal condition (pooled over subjects, lists, and sessions) and Fig. 4 shows the results for the controlled-rehearsal condition (also pooled over subjects, lists, and sessions).

The separate panels in each figure are conditionalized upon number of rehearsals. The five panels are for 1, 2 to 3, 4 to 6, 7 to 10, and 11+ rehearsals. These cuts were selected because they gave a reasonable approximation to an equal frequency division and seemed to be a fair way to divide up the range. For the data shown in Fig. 3, the frequencies in the five rehearsal conditions were 1608, 4196, 2518, 1039, and 879, whereas for the data shown in

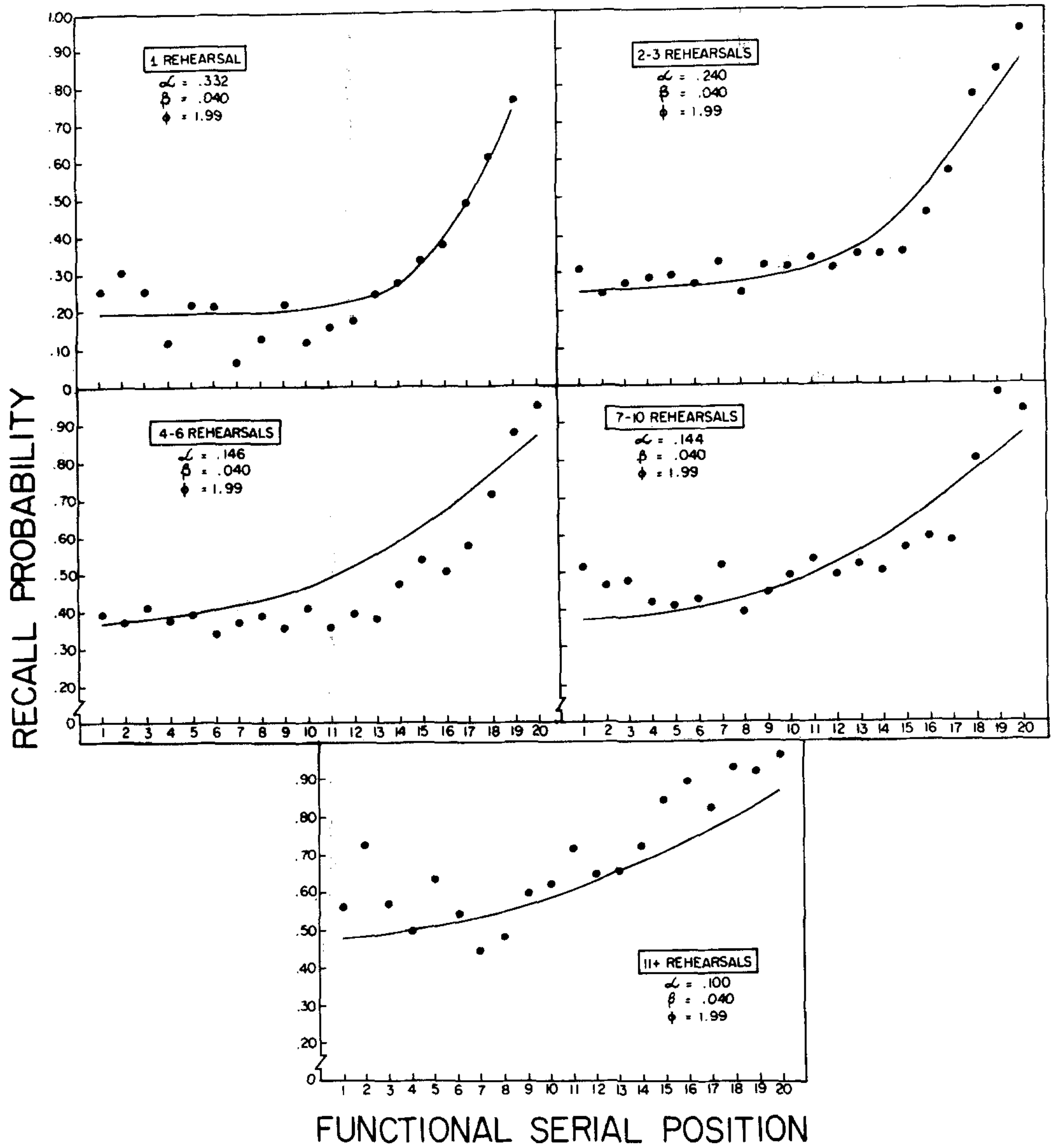

FIG. 4. Functional serial-position curves for the controlled-rehearsal sessions partitioned on the number of rehearsals.

Fig. 4 the corresponding frequencies were 1015, 3867, 3580, 1119, and 659, respectively.

The main reason for partitioning on number of rehearsals is to control for one of the artifacts mentioned in the Introduction. If one of the ways that functional serial-position curves are contaminated by item-selection artifacts is through a confounding of item difficulty and number of rehearsals, then partitioning on this variable will reduce the confounding. It will reduce the confounding in the overt-rehearsal procedure and *magnify* item-selection artifacts. That is, the curve for 11+ rehearsals should be based exclusively on the easiest items, and so recall performance should be very high over all serial positions. The curve for once-rehearsed items should be composed exclusively of the hardest items, so recall performance should be very low. Intermediate rehearsal values should be ordered accordingly. There are, however, no such item-selection artifacts to magnify in the controlled-rehearsal condition. Thus, if the pattern of results is essentially the same, then this analysis would greatly weaken the potential criticism of the functional serial-position analysis.

Obviously the pattern is in fact very similar in the two figures. Therefore, it would seem quite difficult to maintain the item-selection criticism in the face of these results. Furthermore, it is obvious that the number of rehearsals is in itself a very potent variable. This is obvious by inspection, but as further proof consider the probability of a correct recall summed over serial position. For the overt-rehearsal condition the probabilities were .301, .399, .539, .721, and .914 for 1, 2 to 3, 4 to 6, 7 to 10, and 11+ rehearsals. For the controlled-rehearsal condition, these probabilities were .286, .414, .473, .542, and .737, respectively. These data would seem to be quite consistent with the arguments of Nelson (1977) in his recent critique of the "levels-of-processing" view of memory.

In partitioning on rehearsals, there is the danger of a confounding with subject ability. If subjects who rehearse more recall more, then the curves shown in Figs. 3 and 4 might be telling us more about subjects of differential ability than about items of varying strength. This danger seems more apparent than real. There was no correlation across subjects between amount of rehearsal and recall. More specifically, we found the total number of rehearsals and the total words recalled for each subject for the two overt-rehearsal sessions. The correlation between rehearsal and recall was a nonsignificant −.184.

In fitting the Lockhart attribute model to the data shown in Figs. 3 and 4, it was not clear initially how best to represent the role of repetition in the model. Three variations were tried, letting α, β, or ϕ vary with repetition (the other two were not allowed to vary). The best

TABLE 2

PARAMETER ESTIMATES AND GOODNESS OF FIT FOR OVERT REHEARSAL AND CONTROLLED REHEARSAL

	Overt	Controlled
α_1	.450	.332
α_{2-3}	.278	.240
α_{4-6}	.180	.146
α_{7-10}	.122	.144
α_{n11+}	.066	.100
β	.025	.040
ϕ	3.12	1.99
χ^2 (92)	210	272

fit was achieved with α free to vary with repetition, whereas β and ϕ were constant. These are the parameter values shown in Figs. 3 and 4. A summary of this analysis is given in Table 2, which shows not only the parameter values but also the resulting goodness of fit (χ^2). While the fit is not perfect, it seems reasonably good considering the complexity of the processes which generate the data and the relative simplicity of the model which attempts to explain them.

The final free recall curves for nominal and functional serial position are shown in Fig. 5. As before, the nominal curve is based on presentation position while the functional curve is based on rehearsal position. The nominal serial-position curve shows the usual results; namely, a primacy effect, slight

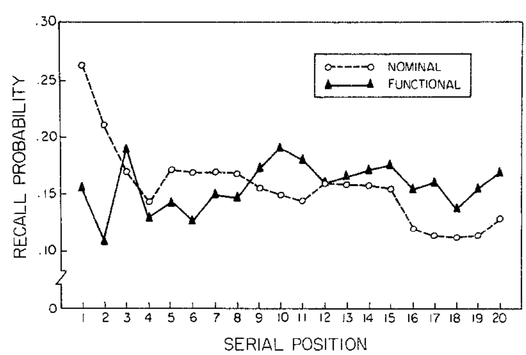

FIG. 5. Nominal and functional serial-position curves for final free recall for the controlled rehearsal sessions.

negative recency, and, by and large, a monotonic decrease from first to last serial position. In contrast, the functional serial-position curve is quite flat; if anything, it is slightly increasing rather than decreasing with serial position.

Two ANOVAs were done, one for nominal and one for functional serial position. The basic classification was subjects (16), sessions (2), lists (16), and serial position (20). Subjects and lists were each pooled into four blocks of four, so the macroblock size was 16 and E(MS) = .063. Sessions were not significant, $F < 1.0$, but lists were, $F(3, \infty) = 50.6$, $p < .001$. In the nominal analysis, serial position was highly significant, $F(19, \infty) = 4.43$, $p < .001$, but in the functional analysis it was only marginally significant, $F(19, \infty) = 1.81$, $p < .05$. Discussion of the effect of the lists will be postponed to a discussion of the results of the Session 6 data, where the basic finding was replicated and extended.

Perhaps of more interest than the serial-position effects themselves are the final recall data partitioned on number of rehearsals, as in Figs. 3 and 4. Mean recall probabilities in final free recall for items with 1, 2, to 3, 4 to 6, 7 to 10, and 11+ rehearsals were .092, .141, .165, .199, and .235, respectively. Clearly, the rehearsal effect is not limited to immediate free recall. It shows up in final free recall as well. This result also holds when one partitions the number of rehearsals into high, medium, and low for each subject separately.

The results so far seem to show relatively little difference between the overt-rehearsal condition and the controlled-rehearsal condition. Now, let us consider the Session 5 data. In that session, the Session 1 lists were represented under controlled-rehearsal conditions with the exact presentation order either preserved or scrambled. Mean recall proportions were .521 and .532 for the preserved and the scrambled lists, respectively. An analysis of variance gave F values of essential 1.0 for both the condition (i.e., type of presentation) and the condition by serial position interaction. Thus, scrambling the rehearsal items the subject had voluntarily selected on a particular list had no detectable effect on his recall performance.

SUPPLEMENTARY EXPERIMENTS

In addition to the main experiment, two supplementary experiments were run to tidy up some loose ends. Since the uncertainties involved points of interpretation of the data from the main experiment, it seemed best to try to clarify these issues by retesting the same subjects. Hence the two supplementary experiments can be referred to as Session 6 and Session 7. Potential criticism about confoundings with practice effects can be allayed partly by reference to the result already reported that such effects were negligible and partly by the fact that, in Session 7, the critical comparison was within sessions rather than between sessions.

Session 6

One of the potential criticisms of the overt-rehearsal procedure is that the rehearsal requirement itself changes the nature of the task. Whatever the subject is doing under overt rehearsal is probably different from what he would be doing without the rehearsal requirement, so generalization from one to the other is risky. The same criticism might well be leveled at our controlled-rehearsal procedure. The subject, left to his own devices, might well allot his attention differently than when faced with the rehearsal words we show on the screen and saying them aloud.

There is already some evidence that this is not a serious problem. Horton (1976) compared overt-rehearsal conditions with silent-rehearsal conditions on three measures: overall recall, proportion of related and unrelated items recalled, and clustering of related items. None of these three measures showed a significant effect due to type of rehearsal. However, the problem seems sufficiently important so some replication would be desirable. That was essentially the point of Session 6.

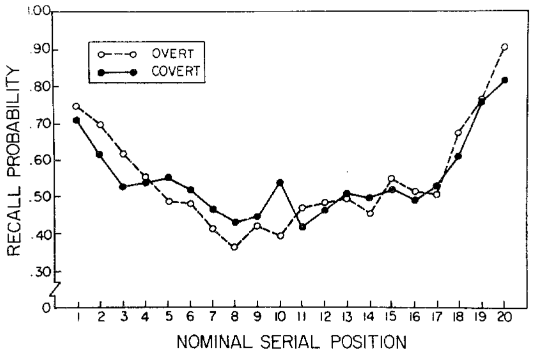

FIG. 6. Nominal serial-position curves for those subjects who participated in both the overt- and covert-rehearsal sessions.

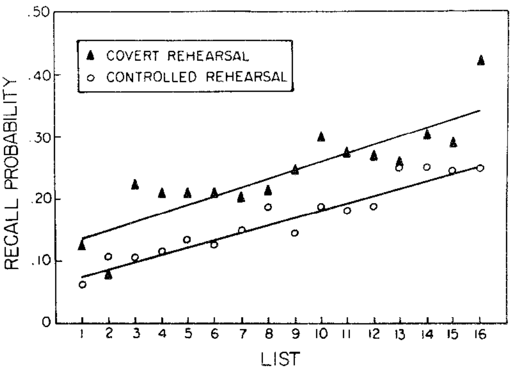

FIG. 7. Final free recall as a function of presentation list for those subjects who participated in both the covert- and the controlled-rehearsal sessions.

Method. Presentation conditions here were exactly like those of the overt-rehearsal condition, but there was no rehearsal requirement. Subjects were simply told to do whatever they wished. We were able to get 10 of the original 16 subjects to return. Except for the rehearsal requirement, all the details of the testing, the list construction, etc. were the same as in the overt-rehearsal sessions. A final recall test was also included.

Results. The main results are shown in Fig. 6, which gives the nominal serial-position curves of these 10 subjects for Session 6 (covert) and for their prior overt-rehearsal condition (Session 2 or Session 4, depending upon counterbalancing group). Clearly, there is little or no difference between the overt-rehearsal and the covert-rehearsal conditions. Summed over serial positions, the mean proportion recalled was .550 for the overt-rehearsal session and .547 for the covert-rehearsal session. Thus, as Horton (1976) found, we also find little tangible evidence to sustain the criticism that the rehearsal requirement fundamentally changes the nature of the task. (To be somewhat more precise, by definition changing the rehearsal requirement changes the nature of the task; the point obviously is that the data are quite impervious to these changes.)

The final recall serial-position curve was much like the nominal curve shown in Fig. 5. Of greater interest is the list effect, previously mentioned for the controlled-rehearsal data. The list effect seems to be linear, and the data for Session 6 as well as the controlled-rehearsal data for the same 10 subjects are shown in Fig. 7. (The data for all 16 subjects of the controlled-rehearsal sessions were much the same as this subset of 10 subjects.) The linear functions shown in Fig. 7 are the least-squares fits; slopes were .012 and .014; intercepts were .065 and .122; and r^2 values were .920 and .769 for the controlled- and the covert-rehearsal sessions, respectively. Clearly, the two conditions differ in the intercept of this function but not the slope. This result was quite unexpected, and we are not quite sure what it means. It seems, however, to be a provocative finding and perhaps worthy of additional investigation.

Session 7

We have essentially demonstrated so far that the overt-rehearsal procedure, with its (not-so-serious) confoundings, is not necessary to generate the functional serial-position effects previously reported. The controlled-rehearsal procedure, without selection artifacts, yields essentially the same results. What produces these results? The conclusion we would like to draw is that the rehearsal pattern is critical, since this was constant in both conditions (i.e., overt and controlled rehearsal).

This would be, however, a conclusion reached by default. While we have controlled for item-selection artifacts, we have not actually manipulated the rehearsal pattern. Such manipulation is necessary to establish the suggested conclusion, and this was the purpose of Session 7.

Method. We used the controlled-rehearsal method on the computer for this session with three different rehearsal conditions. These conditions varied only in Distribution (b), the lag distribution used to select the to-be-rehearsed items. In the normal condition, each subject had the same distribution used in the previous controlled-rehearsal condition. In the uniform condition, we used a uniform (rectangular) distribution so that Word 1, 2, 3, ..., i had an equal chance of being selected as a rehearsal item at Serial Position i, $i = 1, 20$. In the reversed condition, the distribution from the normal condition was simply reversed front to back. That is, at Serial Position i the probability formerly assigned to i was assigned to Serial Position 1, i-1 to Serial Position 2, i-2 to Serial Position 3, etc. In effect, this meant that at every serial position the most likely rehearsal word was the first word in the list, the next most likely the second word in the list, and so on.

The reasoning behind this manipulation was as follows. Without an adequate theoretical model of the rehearsal process, it is hard to know how to improve on whatever rehearsal pattern the subject himself adopts. One should be able, however, to decrease its effectiveness, so that was the approach tried here. The expectation was that performance would be best under the normal condition, perhaps somewhat depressed under the uniform condition, and considerably depressed under the reversed condition.

Actually, the argument can be made stronger than this. The main reason we should get differences is that the rehearsal effects are nonlinear. A few items, particularly for the reversed condition, are receiving a disproportionate number of rehearsals. It is not helping them much, but it is detracting from other items. Since recall is presumably a function of item strength, and since total strength is the sum of the item strengths, what one gains by greatly strengthening a few items is more than offset by loss on the remainder. Exactly how this should work requires an explicit model, but at least the general principle seems clear enough.

We were able to get 7 subjects of the original 16 to return; all of them had participated in Session 6, too. For each of these subjects, there were 15 lists in Session 7, 5 of each kind. The order of conditions within the session was randomized. Presentation conditions and details of the experimental procedure were the same as in the two earlier controlled-rehearsal sessions. It should be noted that both the rationale and the experimental manipulations used in Session 7 were very similar to those of Experiment II in Brodie and Prytulak (1975). Essentially, what they did with overt rehearsal we did with controlled rehearsal.

Results. The mean recall proportions were .507, .487, and .396 for the normal, uniform, and reversed conditions, respectively. An analysis of variance, pooled over the five lists per condition, showed that the effect was highly significant; $F(2, \infty) = 10.74$, $p < .001$. Thus, the predicted result was obtained; changing the rehearsal pattern did depress performance, and the more drastic change (the reversed condition) had a larger effect than going from the subject's own rehearsal pattern to a uniform distribution.

Figures 8 and 9 show the nominal and functional serial-position curves for the three conditions of Session 7. As expected, the last few items are better recalled in the normal condition, while the first few items are better recalled in the reversed condition. These recency and primacy effects extended over only a few serial positions; these results could be expected because the lag distributions were very sharp (see Table 1). The functional serial-position curves seem to show little or no

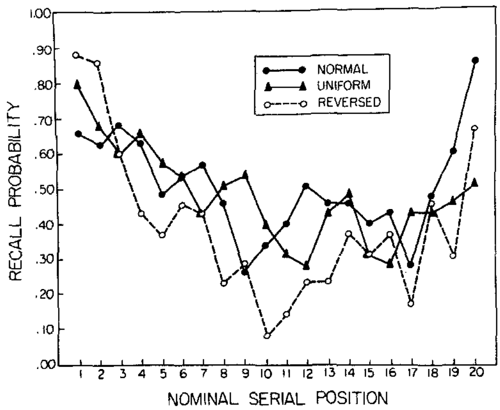

FIG. 8. Nominal serial-position curves for the three conditions of Session 7.

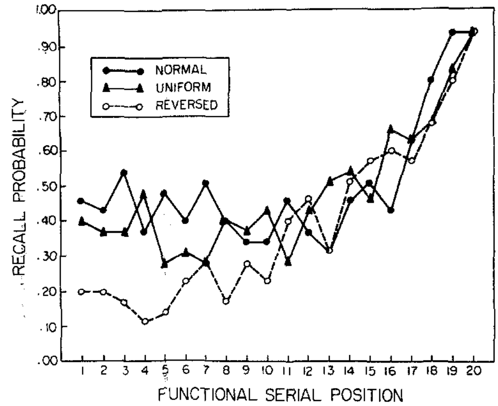

FIG. 9. Functional serial-position curves for the three conditions of Session 7.

difference over the last half (i.e., Positions 11–20) but fairly regular differences (normal better than uniform better than reversed) over the first half (Positions 1–10). This would suggest that the last rehearsal position overrides the number and distribution of rehearsals for recent items, but the effect of these variables does show up for early items. At least in this one case, the functional serial-position curves would seem to provide more insights about the underlying processes than the nominal serial-position curves. Also, these two curves provide a dramatic example that primacy effects depend upon how the data are analyzed.

DISCUSSION

These data have increased the credibility of the overt-rehearsal procedure for single-trial free recall. Any stigma due to item-selection artifacts has been removed. Though item-selection effects can occur in the overt-rehearsal procedure, we find similar results in a controlled-rehearsal procedure where item-selection effects cannot occur.

This is not to say that there are no differences at all between these two procedures. There was slightly better recall under the overt-rehearsal procedure, and the parameters of the Lockhart attribute model were somewhat different. The Session 5 data would seem to preclude item-selection effects as being the explanation. Another difference between the two procedures is that the controlled-rehearsal is more "passive" than the overt-rehearsal procedure. One simply watches the display, and the work is all done. Perhaps the effort of retrieving items to rehearse them is what produces the difference. In the overt-rehearsal condition one must retrieve each item to rehearse it, so each rehearsal functions as a practice retrieval. There is much evidence (see, e.g., Murdock, 1974, Pp. 245–249) that repeated retrievals (test trials) facilitate recall over trials. Perhaps the same effect occurs within a single trial. Such an effect would be over and above that accruing from sheer repetition.

Greater differences between overt and controlled rehearsal might show up in multitrial free recall. If nothing else, one would think that the overt-rehearsal procedure would be more beneficial for chunking than the controlled-rehearsal procedure. Of course, this really leads to the question of the effect of presentation order on learning, and Murdock, Anderson, and Ho (1974) found surprisingly little effect. In their Experiment II, clusters which were preserved in list presentation were not learned much better than clusters which were scrambled. The serial-position curves of Session 7, along with the general pattern of data, suggest a possible answer. Given the same total presentation time, about the same number of words will be recalled regardless of presentation order. Rehearsal patterns will determine which words will be recalled but not, to any large extent, how many. Only with gross distortions (e.g., the reversed condition of Session 7) will total number recalled be affected. The functional serial-position curves are more invariant over conditions than the nominal serial-position curves.

How can we explain the serial-position data? The Lockhart attribute model describes the functional curves, but that is not enough. A comprehensive model must explain both functional and nominal serial-position curves. While nominal and functional curves are quite different, they are different representations of the same data. While there are some arguments for preferring functional to nominal analyses, much of the empirical and theoretical literature is based on nominal serial-position curves. Consequently, the nominal and functional analyses should be derivable from common principles.

Here is a general framework for providing the desired integration. Let P be the nominal–functional position matrix, the joint probability that a given item is in Nominal Position i and Functional Position j in a list of length N. This matrix P can be obtained from the overt rehearsal data or could be derived from more basic assumptions about the nature of the rehearsal process. Let r be the N-element (column) rehearsal vector, the number of rehearsals at each nominal serial position. Let d be the decay vector, a row vector for the strength of an item as a function of the number of other items rehearsed subsequently. The vector d is based on functional serial position. Then $S = r \times d$, an $N \times N$ strength matrix indicating the hypothetical strength of an item at Nominal Position i and Functional Position j.

If there is a direct relationship between strength and recall probability, then the functional serial-position curve can be derived from the diagonal elements of the matrix $S'P$ (S' denoting the transpose of S) and the nominal serial-position curve can be derived from the diagonal elements of the matrix PS'. The diagonal element of $S'P$ form a vector which, when normalized to the number of words recalled, gives a predicted functional serial-position curve. Likewise, the diagonal elements of PS' form a vector which, when normalized to the number of words recalled, gives a predicted nominal serial-position curve. Thus, both nominal and functional serial-position curves can be derived from very simple assumptions about the underlying rehearsal processes.

The decay vector d needs further specification. As it now stands, it simply describes the strength value of each item following its last rehearsal. However, this is the end result, and it would be nice to include the effect of the number and distribution of prior rehearsals. Here is one attempt. Assume each rehearsal of an item increases its strength by one unit and decreases the strength of all other items by k%. At the end of list presentation, the stronger items are recalled while the weaker items are not recalled. A value of k is desired which maximizes the difference between the two distributions (i.e., the strength distribution of the recalled items and the strength distribution of the nonrecalled items).[2]

We ran a parameter estimation of this model for its possible heuristic value using the data of each subject on the overt-rehearsal lists and the controlled-rehearsal lists. There were 32 lists of each type for each subject. For each list, the program followed the rehearsal protocol, incremented and decremented the strength of each item with each presentation or rehearsal, and estimated the best value of k with SIMPLEX. The criterion for "best" was that value of k which gave the largest point

[2] This approach was suggested to us by Howard Kaplan.

biserial correlation, the appropriate statistic when one is trying to relate a continuous variable (strength) with a discrete variable (recall or nonrecall of each item) (McNemar, 1962). (It is interesting to note the similarity between this measure of correlation and d', the standard strength measure of signal-detection theory.) The parameter estimates (i.e., k) ranged from about 1 to 20% across subjects, and the goodness of fit (i.e., the values of the point biserial correlation) was reasonably high. The major disappointment came when we compared the parameter estimates for the two conditions (overt rehearsal and controlled rehearsal). The correlation (across subjects) was essentially zero. One would think there should be a correlation, and its absence is puzzling.

One of the more interesting aspects of this model was the strength distributions at different points during the presentation of the list. While the principles in general are clear enough (increment an item by 1, decrement all others by k%), how this works out in detail is not immediately obvious. To find out, we had the program print out the hypothetical strength values for each item at various points in the list. After a few items had been presented, there was a marked primacy effect. The first item was strongest, and later items were progressively weaker. As list presentation continued, the marked primacy effect gradually evaporated and a recency effect began to emerge. By the end of list presentation the primacy effect was almost absent, but the recency gradient was very prominent. Such a view has interesting implications both for serial recall (where short lists are used, and primacy greatly outweighs recency) and for the change from primacy to recency that seems characteristic of practiced subjects in single-trial free recall.

This rehearsal model is really only a rough outline, but it does have some virtues. It shows how nominal and functional serial-position curves can both be derived, quite simply in fact, from assumptions about underlying processes. It emphasizes traditional strength and interference factors, much as suggested by Brodie (1975), Brodie and Murdock (1977), and Brodie and Prytulak (1975). It emphasizes rehearsal as a dynamic process, with the resulting interactions among items. These interactions are not easy to characterize, but it would seem a serious oversight to neglect them. The only other model we know of that has attempted this has been suggested by Bernbach (1969), and his rehearsal model is similar in several respects.

Where does this leave the Lockhart attribute model? Frankly, we are not quite sure. One possibility is that it is simply a useful empirical description but little more. That is, it certainly does characterize quite adequately the functional serial-position curves, both for the overt- and the controlled-rehearsal procedure. Consequently, one can describe them as three-parameter functions, these parameters being ϕ, α, and β. These parameters provide a very detailed account of the functional serial-position curves. With them, one can generate (compute) all points on the curve. This would seem to be a more analytic account than the traditional Primary and Secondary Memory measures of nominal serial position. However, it may be that the Lockhart model has a theoretical contribution to play as well. Perhaps it could describe the decay process, only the decay process would apply at the level of attributes but not items. Some further theoretical effort will be necessary to decide.

Finally, the experimental implications of this research should be mentioned. Controlled rehearsal can now be used as a supplement to or in place of overt rehearsal. It greatly simplifies the experimental effort, since scoring the rehearsal protocols is now replaced by recording the presentation order. More important, explicit manipulations of rehearsal patterns are now possible. If rehearsal really is the important variable as many of us have long assumed then controlled rehearsal, along with comparisons of nominal and functional serial-position curves, may help us gain better understanding.

References

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1968. Vol. 2, pp. 89–195.

Bernbach, H. A. Replication processes in human memory and learning. In G. H. Bower and J. T. Spence (eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1969. Vol. 3, pp. 201–239.

Brodie, D. A. Free recall measures of short-term store: Are rehearsal and order of recall data necessary? *Memory & Cognition*, 1975, **3**, 653–662.

Brodie, D. A., & Murdock, B. B., Jr. Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 185–200.

Brodie, D. A., & Prytulak, L. S. Free recall curves: Nothing but rehearsing some items more or recalling them sooner? *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 549–563.

Corballis, M. C. Patterns of rehearsal in immediate memory. *British Journal of Psychology*, 1969, **60**, 41–49.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 671–684.

Glanzer, M. Storage mechanisms in recall. In G. H. Bower (ed.), *The Psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1972. Vol. 5, pp. 129–193.

Horton, K. D. Phonemic similarity, overt rehearsal, and short-term store. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 244–251.

McNemar, Q. *Psychological statistics*. New York: Wiley, 1962. 3rd ed.

Murdock, B. B., Jr. *Human memory: Theory and data*. Potomac, Md.: Lawrence Erlbaum Associates, 1974.

Murdock, B. B., Jr., Anderson, R. E., & Ho, E. Effects of presentation order on learning in multi-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, 1974, **13**, 522–529.

Murdock, B. B., Jr., & Ogilvie, J. C. Binomial variability in short-term memory. *Psychological Bulletin*, 1968, **70**, 256–260.

Nelder, J. A., and Mead, R. A simplex method for function minimization. *Computer Journal*, 1965, **7**, 308–313.

Nelson, T. O. Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 151–171.

Rundus, D., & Atkinson, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 99–105.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon and D. L. Horton (eds.), *Verbal Behaviour and General Behavior Theory*. Englewood Cliffs: Prentice–Hall, 1968. Pp. 2–36.

(Received September 12, 1977)