The replacement effect: Repeating some items while replacing others

BENNET MURDOCK and MARY LAMON
University of Toronto, Toronto, Ontario, Canada

We performed three experiments on recognition learning that tested for the existence of a replacement effect (i.e., the benefit accruing to nonrecognized items, or targets, when recognized items are replaced in the next study trial). A reverse Rock substitution procedure was used, and the replacement effect occurred in all three experiments. The results were interpreted in terms of a distributed memory model, the matched-filter model of Anderson (1973), but several modifications were necessary. The original version cannot learn, and a closed-loop modification did not show the repetition effect that was clearly evident in the data. The most satisfactory version was one based on probabilistic encoding of features in the item vectors, and it seemed capable of explaining most aspects of the data.

Recent developments in distributed memory models provide a new way of looking at human memory. They specify how information can be stored and retrieved, they explain how recognition and recall can occur, and, coupled with a decision system, they can account for both accuracy and latency data. They are content addressable, and retrieval occurs by direct access rather than by search. They are consistent with known physiological mechanisms; in fact, many distributed memory models (see Hinton & Anderson, 1981) are based on or are outgrowths from neural models.

Originally, it was not clear how to test distributed memory models, so their usefulness to experimental psychologists was limited. However, a number of applications to experimental data have appeared more recently (e.g., Anderson, 1973; Anderson, Silverstein, Ritz, & Jones, 1977; Eich, 1982, 1985; Heath, 1983; Humphreys, Pike, Bain, & Tehan, in press; Kawamoto & Anderson, 1985; Knapp & Anderson, 1984; Lewandowsky & Hockley, 1987; McDowd & Murdock, 1986; Murdock, 1982, 1983; Nilsson, 1986), so the models can clearly make contact with experimental data.

A further test is whether distributed memory models predict any novel or unexpected empirical phenomena. We argue that they do, and illustrate with a particular example. The prediction to be tested is that in a recognition learning situation, replacing items recognized on one trial should facilitate recognition of the repeated items ("targets") on the next trial. This facilitation is relative to a control condition in which there is no replacement.

More specifically, in the experimental or replacement condition, one replaces all items recognized on the test phase of Trial $n-1$ with a new sample of items for the study phase of Trial n. In the control or nonreplacement condition, all old items from Trial $n-1$ are repeated in the study phase of Trial n. The prediction is that performance on targets (the nonrecognized items) should be better than performance on nontargets in the experimental condition but worse in the control condition. To facilitate exposition, we call this prediction the "replacement effect."

Why do distributed memory models predict the existence of a replacement effect? To understand this prediction, we must first understand why some items are recognized and some items are not recognized on any given trial. Basically, each item is masked by all the other items on the list, and the amount of this masking varies from item to item. Purely by chance, some items are at the bottom of the distribution, and those are the items that are not recognized. As signal-detection theory tells us, memory-trace strength is graded not all-or-none, and items below the criterion have less strength than other items but not zero strength.

What should happen to the target items if one replaces the masking items with a new (random) sample? By replacing the recognized items, there should be a "regression to the mean" for the target items. Although they were below average for the first sample of items, with a new sample of items there is no reason they should still be below average, hence the regression to the mean. As a result, the target items should be better recognized in the replacement condition than in a standard or control nonreplacement condition. (We are oversimplifying at this point, because there are two opposing factors. Targets in the experimental condition are handicapped by being of greater than average difficulty but, compared with the nontargets, are benefited by the repetition. How these two factors trade off will be explicitly considered in the Discussion section.)

The predicted replacement effect highlights some of the differences between distributed memory models and a more standard type of strength theory (e.g., Wickelgren

& Norman, 1966). According to a distributed memory model (e.g., the matched-filter model of Anderson, 1973), there is no discrete representation of individual items; all items are pooled in a common memory vector. Any forgetting that occurs is global, not local; one item cannot be forgotten more than another. Retrieval occurs by a direct comparison of a probe item to the memory vector, and whether or not old items are recognized depends on how much they "resonate" at the time of testing. Distributed memory models are what Humphreys et al. (in press) call "global matching models."

However, it turns out that the picture is not quite so simple. The prediction about the replacement effect was arrived at intuitively and, although it seems quite reasonable, rests on a presupposition that turns out to be false. The presupposition is that distributed memory models can learn, but in fact some of them cannot. Therefore, we must talk about particular distributed memory models, not distributed memory models in general. We will start with the matched-filter model of Anderson (1973), a simple but elegant distributed memory model for item recognition.

The Matched-Filter Model

According to the matched-filter model, the storage and retrieval of item information is as follows: Item information is represented by N-dimensional random vectors. These are vectors in an N-dimensional space whose elements are random variables. Each of the N elements in an item vector is a random sample from a normal distribution with mean zero and variance P/N, where P is the power of the vector. ($P = \sum_{i=1}^{n} f_i f_i$, where N is the number of elements in the vector.) Each item is stored in a common memory vector $\mathbf{M}$. The storage of item information is given by:

$$\mathbf{M}_l = \alpha \mathbf{M}_{l-1} + \mathbf{f}_l, \qquad (1)$$

where the item vectors are represented by $\mathbf{f}_l$ for $l = 1$, $2, \ldots, L$, and L is the number of items in the list. For retrieval, the probe item is compared with the memory vector and the comparison operation is the dot product. This dot product is fed into a decision system whose output is an observable response, "no" or "yes" depending on whether the value of the dot product falls below a lower criterion or above an upper criterion. (The matched-filter model is a two-criterion model.)

A major problem with the matched-filter model is that it cannot account for recognition learning. Memory after a single presentation can be quite good or bad depending on the number of items in the list and the values of the parameters N and α, but it can be just as good (or just as bad) after the second presentation, the third presentation, and so forth. That is, there is no improvement in performance with repetition. We know (e.g., Kintsch, 1965) that this is wrong.

Why does learning not occur with repetition? Each repetition adds a new copy of the item to the memory vector, and this increases the expected value of the dot-product

comparison. However, it also increases the variance. Since d', the standard measure of recognition memory, depends on both the old-item mean and the new- and the old-item variances, both the mean and the variance must be considered.

The fact that both the mean and the variance increase with repetition is not enough to say that learning does not occur. What is necessary is to know what happens to d' with repetition, as d' takes into account both mean and variance. In fact, d' does not change as the number of presentations increases, and a proof is given in Appendix A.

An obvious modification to rectify the learning problem is to change the matched-filter model from an open-loop system to a closed-loop system. In a closed-loop system the output modifies the input, whereas it does not in an open-loop system. This difference is illustrated in Figure 1. The homeostatic regulatory mechanisms of the body are well-known examples of closed-loop systems, and such closed-loop or feedback systems have been much discussed in the literature (see, e.g., Arbib, 1964, or Wooldridge, 1963). Closed-loop or feedback systems have been widely used in neural models of learning (e.g., Sutton & Barto, 1981) and in connectionist models (e.g., McClelland & Rumelhart, 1985). Also, this closed-loop model can account quite well for some recognition-memory data of Avant and Bevan (1968) on constant versus varied encoding (McDowd & Murdock, 1986), so it seemed reasonable to try it here.

A simple way to turn an open-loop system into a closed-loop system is to rewrite Equation 1 so that

$$\mathbf{M}_l = \mathbf{M}_{l-1} + c\,\mathbf{f}_l, \qquad (2)$$

where $c = 1 - d$ and d is the dot product of item l with the memory vector $\mathbf{M}$ (McDowd & Murdock, 1986). We let retrieval precede storage, and we let what is stored be a function of what is retrieved. In other words, assume that when a study item is presented, it is first compared to the information in the usual way (i.e., dotted with the memory vector); then the additional information that is added to the memory vector is inversely proportional to what is already there. When the item is familiar (i.e., when d,

OPEN - LOOP SYSTEM

INPUT ⟶ [ ] ⟶ OUTPUT

CLOSED - LOOP SYSTEM

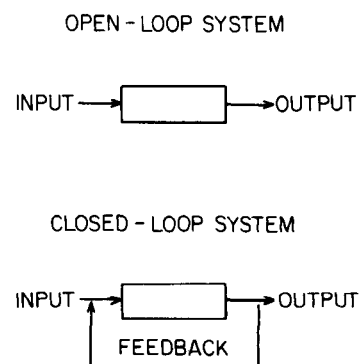INPUT ⟶ [ ] ⟶ OUTPUT
        FEEDBACK

Figure 1. Illustration of the difference between an open-loop and a closed-loop control system.

the dot product, is large) little is added, but when the item is unfamiliar (i.e., when d is small) much is added. The constant c acts as a negative-feedback weighting to determine how much of the item is added to memory.

As we show here, this closed-loop version does indeed rectify the learning problem, so our presupposition for the replacement effect is well-founded. Whether the closed-loop version can explain the replacement effect is another matter. We turn now to the experiments that were designed to provide the data to test the model.

METHOD

The experiments were designed to test whether or not the replacement effect occurs, and they used what may be called a reverse Rock substitution procedure (RRSP) (Lamon, 1982). In the experimental condition all items recognized on the test phase of Trial $n-1$ were replaced for the study phase of Trial n. This procedure contrasts with the original substitution procedure of Rock (1957), in which all nonrecalled pairs were replaced. A similar procedure has been used in paired-associate learning (e.g., Cavanagh & Parkman, 1971), but not, to our knowledge, in recognition.

Using the RRSP, we focused on the recognition of the target items on the test phase of Trial n. The prediction was that performance on the target items should be better than performance on the nontarget items in the experimental condition but not in the control condition.

We conducted three separate experiments, in part to vary the procedural details in order to test the generality of the replacement effect. Also, we varied the difficulty level of the task to see if the replacement effect interacted with task difficulty. We detail the procedure and subjects for each experiment separately, and then discuss the results.

Experiment 1

A discrete-trials study–test recognition procedure was used. Lists of 80 words were presented singly to each subject. After study-list presentation, the subject was given a 160 word yes–no recognition test. The test consisted of the 80 study words and an equal number of new words that served as lures. In the experimental condition, correctly recalled items on Trial $n-1$ were replaced for the study trial on Trial n. If a study item was not recognized, it was repeated on the next study list. These repeated items were the target items, and target items were defined in terms of only the previous trial. They were missed on Trial $n-1$, and hence repeated on Trial n. What happened on Trial $n-2$ was irrelevant.

In the control condition, the same 80 study items were presented and tested with new lures on each test trial. There were target items (the nonrecognized items on the prior trial), but the recognized items were not replaced. Three study–test trials were given for each condition. Only three trials were used because preliminary testing showed that performance was near the ceiling after three trials.

The subjects were assigned randomly to conditions, and each subject was tested individually. The subjects were asked to try to learn each word as well as possible and were informed that a recognition test would follow immediately. The subjects were told that on subsequent study lists some proportion of words studied on the first list would be repeated, but that the task did not involve discriminating among lists. Study cards, each bearing a single typed word, were presented one at a time at a rate of 30 words per minute. The recognition test followed immediately and consisted of items shown on cards presented one at a time. The recognition test was self-paced. The subjects were asked to respond to each item, indicating whether or not they recognized the word as a study word by placing the test card on a 2×2 matrix. Old items were to be placed to the left

and new items to the right. Each test card was coded on the reverse so that the experimenter knew whether that item was old or new. The experimenter then indicated whether the card should be placed in the upper (correct responses) or lower (incorrect responses) quadrant. The subjects were informed about this method of scoring; thus, they had immediate feedback about accuracy of performance for each item.

A word pool was constructed by selecting common nouns that had at least 30 occurrences per million words according to the Kučera and Francis norms (1967). These words were matched for length and concreteness. A noun was classified as concrete if it met either of two criteria: (1) it was rated as greater than 6.0 on the Paivio, Yuille, and Madigan norms (1968), or (2) if there was no available rating, the word had to be a reasonably appropriate object of a sense verb, such as see, hear, touch, taste, or smell (Glanzer & Bowles, 1976). There were approximately 500 words in the resulting word pool that were assigned randomly without replacement as targets or lures in constructing unique sets of study and test lists for each subject. For each control subject, a single study list and three sets of lures were constructed. For each experimental subject three old-item sets and three sets of lures, each consisting of 80 items, were constructed. Order of item presentation was randomized for each study and test list.

The subjects were 30 volunteers, male and female, ranging in age from 18 to 24 years. The subjects volunteered in order to receive course credit. A 3×2 split-plot design was used. Repetition was a within-subjects factor and replacement condition was a between-subjects factor.

Experiment 2

The same RRSP was again used for the replacement manipulation. This experiment was performed on an IBM PC. The computer was used both to generate study and test lists and to administer the experiment. The subjects were presented with a list of 64 words displayed one at a time in the center of a CRT screen for 1.25 sec with an interstimulus interval (ISI) of 300 msec. The subjects were informed that immediately following list presentation a recognition test list would be presented, with each word appearing alone. The subject's task was to respond to each item as it appeared by pressing a "yes" key if the item was recognized as a study word or a "no" key if it was not recognized as a study word.

Response keys were assigned randomly so that half of the subjects made "yes" responses with their preferred hand and half made "yes" responses with their nonpreferred hand. When ready, the subject pressed the "yes" key to begin testing. The recognition test was self-paced. Latencies were recorded, and as soon as the subject made a response, immediate feedback about accuracy and latency was given. Either "correct" or "incorrect" appeared below the test item, along with the latency in milliseconds. There was a 300-msec interval between response and display of the next test item. Three study–test trials were given in this manner. The subjects were assigned randomly and blindly to conditions. Each subject was tested individually.

Study and test lists were constructed for each subject by selecting words randomly and without replacement from the Toronto Word Pool. This pool consists of 1,080 common English words from different parts of speech ranging in length from four to eight letters. We chose these words of more general nature to help ensure that the effect found in Experiment 1 was not due to idiosyncratic properties of common concrete nouns.

The subjects were 20 male and female volunteers ranging in age from 17 to 30 years. The subjects were paid for their participation. The design was the same as that used for Experiment 1.

Experiment 3

To increase the difficulty level, Experiment 3 used a longer list and a faster presentation rate. In addition, there was a 5-min inter-

polated distractor task between each trial. The same general design was employed in this experiment as in the first two.

Generally, the same RRSP used in the previous experiments was followed in this experiment. Presentation time was decreased from 1.25 sec per item to 300 msec per item with a shortened ISI of 100 msec. Thus, each word in this experiment was allowed only about one quarter of the time allowed in Experiment 2. List length was also increased to 96 items from 80 items in Experiment 1 and 64 items in Experiment 2. Between each study and test trial and between each test trial and subsequent study trial, a 5-min filled retention interval ensued. During this time the subjects were required to read a prose passage. The subjects were informed that a recall test for the gist of the passage would be given after the third recognition test. Again, the subjects were assigned randomly and blindly to conditions, and each subject was tested individually.

As in Experiment 2, the Toronto Word Pool was used. The same method of list construction and randomization was again employed. A split-plot design was used with trials as a within-subjects factor and conditions as a between-subjects factor. The subjects were 26 male and female volunteers, aged 18-55. Half of the subjects participated for monetary payment; the remainder were undergraduates who volunteered in order to fulfill a course requirement.

RESULTS

The main results of the experiments are shown in Table 1. The table shows hit rates and false-alarm rates for targets and nontargets for the control and experimental groups for all three experiments. All items were targets on Trial 1, so the number of targets on Trial 2 did not differ except by chance for the control and experimental groups. Thereafter, the number of targets differed since the experimental group had a partly new list on each trial and the control group saw the same items repeatedly.

Table 1
Hit Rate (HR) and False-Alarm Rate (FAR) for Targets (T) and Nontargets (N) by Condition and Trials for Experiments 1-3

Trial	Type	Control HR	Control FAR	Experimental HR	Experimental FAR
		Experiment 1			
1	T	.76	.10	.77	.10
2	T	.95		.99	
			.05		.09
	N	.97		.82	
3	T	.96		1.00	
			.03		.10
	N	.97		.80	
		Experiment 2			
1	T	.74	.16	.78	.15
2	T	.89		.96	
			.11		.19
	N	.96		.86	
3	T	.96		1.00	
			.08		.18
	N	.97		.85	
		Experiment 3			
1	T	.59	.31	.60	.31
2	T	.73		.82	
			.23		.32
	N	.86		.59	
3	T	.78		.90	
			.16		.35
	N	.92		.60	

There are a variety of measures we could have used to test for the existence of a replacement effect, but we considered only two. The first and perhaps more obvious measure is the d' value for targets. With this measure one would compare d' for targets in the experimental condition with d' for targets in the control condition on each trial. The prediction, of course, is that d' for targets in the experimental condition should be higher than d' for targets in the control condition. By conventional wisdom, this comparison would correct for criterion effects and, as the false-alarm rates show, such criterion differences seem to exist.

There are two problems with this measure. The first problem is that an erroneous interpretation might result if there were differences in the variance. The use of d' as an experimental measure assumes equal variances, yet our theoretical understanding of this situation suggests that the variances are probably unequal. We have no way of assessing from the data what the underlying variances might be, so although this problem is only hypothetical it is still a worrisome point.

The other problem is that the d' measure focuses on a subset of the data (targets only) and neglects the remainder (i.e., nontargets). To test the replacement effect, we need a measure that reflects performance on targets and nontargets in both control and experimental condition. The hit-rate crossover is such a measure. By "hit-rate crossover," we mean the difference between targets and nontargets in the experimental condition compared with the same difference in the control condition. Symbolically we can represent the hit-rate crossover, H, as $H = (a-b)-(c-d)$ where a is the hit rate for targets in the experimental condition, b is the hit rate for nontargets in the experimental condition, c is the hit rate for targets in the control condition, and d is the hit rate for nontargets in the control condition.

This measure H is like an interaction but since, strictly speaking, it is not an interaction, we use the term "crossover" instead. That is, H is not an interaction in the analysis of variance sense, because item type (target vs. nontarget) is not crossed with the replacement manipulation.[1] However, H is the theoretically relevant measure that the experiments were designed to test. Furthermore, it turns out to be a crucial measure to use to compare the three experiments.

As a statistical test we used a simple t test for the difference between the experimental and control groups, where the data for each group was the target-nontarget difference for each subject. That is, for each experiment we computed $a-b$ for each subject in the experimental group and $c-d$ for each subject in the control group and performed a between-groups t test on these difference scores. Thus, we have a single measure (H) and an appropriate statistical test (a t test) to represent and test what we are calling the hit-rate crossover.

To provide additional documentation we also considered two sets of binary contrasts. One binary contrast is the comparison of a with c, that is, hit rate for experimental

targets with hit rate for control targets. This gives a measure of the effectiveness of context change on unrecognized items, which we call the *context effect*. The other binary contrast is the comparison of *a* with *b*, that is, targets in the experimental condition with nontargets in the experimental condition. This gives a measure of the effectiveness of repetition, which we call the *repetition effect*.

The hit-rate crossover on Trial 2 is summarized in Table 2. The table shows the hit rate for targets and nontargets in both the experimental and control conditions. As can be seen, in all three experiments, targets were worse than nontargets in the control condition but (much) better in the experimental condition. The crossover was significant in all three experiments [$t(18) \geq 4.2$, $t(18) = 3.81$, and $t(24) = 6.36$ for Experiments 1, 2, and 3, respectively; in all cases $p < .01$].[2]

The context effect involves a comparison of the hit rate for targets in the experimental condition with the hit rate for targets in the control condition. In Experiments 1 and 2 there were ceiling effects (about half the subjects in the experimental condition correctly recognized all the targets on the second trial), so a statistical test would not be too meaningful. In Experiment 3 this was not a problem, and, although the difference was in the right direction (see Table 2), it was not statistically significant [$t(24) = 1.41$, $p > .05$].

Consequently, the data do not demonstrate a context effect, but the test was not too powerful. Also, in addition to probable criterion differences (which go the other way), an obvious confounding should be noted. The nontargets in the control group were repeated items the subjects had previously recognized, so the control subjects could have devoted more attention to the target items or rehearsed them more than did the experimental subjects, for whom the nontargets were items they were seeing for the first time. So, although we cannot claim there is a context effect, we would certainly not deny the possibility that one could be found under the right conditions.

The repetition effect was clearly evident and statistically significant in all three experiments. In the experimental group, targets were better recognized than non-

targets, with $t(14) = 6.05$, $t(9) = 3.88$, and $t(12) = 6.09$ in Experiments 1, 2, and 3, respectively (in all cases $p < .01$). In fact, for only 1 subject in all three experiments was the difference in the opposite direction. Thus, the benefit of repetition clearly outweighed the disadvantage of item difficulty, and this was true in all three experiments.

GENERAL DISCUSSION

The results of these experiments demonstrate the existence of a replacement effect, and our chief evidence is the hit-rate crossover. It is a large effect that clearly occurred in all three experiments. The context effect was not statistically significant, but a possible confounding with differential attention or rehearsal was noted. The repetition effect was quite large and statistically significant in all three experiments. Target items were recognized better than nontarget items by the experimental group in all three experiments.

The repetition effect cannot reasonably be attributed to criterion differences. Each subject must set a criterion, but the criterion is between new and old items, not between targets (items with two previous presentations) and nontargets (items with one previous presentation). Both targets and nontargets should be given the same response (i.e., "old"), and to claim that there are different criteria for targets and nontargets puts one in the position of claiming that subjects somehow know the difference between once-presented and twice-presented items without knowing the difference between old and new items. This does not seem to be a very reasonable position to maintain.

The purpose of conducting these experiments was to provide some data that could be used to test a prediction from distributed memory models in general and from the matched-filter model in particular. We now consider whether the model can accommodate these data. As we shall show, the closed-loop version of the matched-filter model cannot; in particular, it cannot generate the repetition effect that was so evident in the data. After we have documented this claim, we shall end by briefly presenting a different modification of the matched-filter model that can learn, generate a repetition effect, and generate a hit-rate crossover that at least qualitatively matches the data from the three experiments.

Before we do this, however, it is necessary to rule out a simpler explanation of the data based on strength theory. It could be argued that any theory that locates item difficulty in the item would predict a hit-rate crossover. Control nontargets are stronger than control targets which are equal to experimental targets which are stronger than experimental nontargets, or $d > (c = a) > b$ (see above). Then, the hit-rate crossover is given by $(a-b) - (c-d) = -b+d = d-b$, which will be positive because $d > b$.[3]

Although this is true, it turns out that this simpler explanation makes the wrong prediction about how the hit-rate crossover should vary with task difficulty. It predicts

Table 2
The Hit-Rate Crossover on Trial 2 for Experiments 1-3

Condition	Targets		Nontargets	
Experiment 1				
Control	.95	(288)	.97	(912)
Experimental	.99	(276)	.82	(924)
Experiment 2				
Control	.89	(166)	.96	(474)
Experimental	.96	(141)	.86	(499)
Experiment 3				
Control	.73	(512)	.86	(736)
Experimental	.82	(499)	.59	(749)

Note—The numbers in parentheses are the frequencies on which the hit rates are based.

that the magnitude of the crossover should decrease as task difficulty increases (see Appendix B). The relevant data are shown in Table 3, which shows the value of d' on Trial 1 and the magnitude of the hit-rate crossover on Trial 2 for each of the three experiments. Since Experiment 3 clearly had the most difficult task yet had the largest hit-rate crossover, this simpler explanation does not seem to be correct.

This analysis makes explicit the trade-off between item difficulty and repetition benefits mentioned in the introduction. Obviously, the benefits of repetition outweigh the item difficulty factor. Furthermore, this analysis highlights the advantage of using the hit-rate crossover as the main dependent variable. Under the linearity assumptions of strength theory, computation of the three-way effect (i.e., the hit-rate crossover × task difficulty) cancels out several terms that might make interpretation difficult (see Appendix B).

We now turn to a more realistic model, the closed-loop version of the matched-filter model, to see if it can handle the replacement effect.

Simulations of the Closed-Loop Model

Unlike the open-loop version, it is difficult, if not impossible, to obtain explicit expressions for the closed-loop version of the matched-filter model. Consequently, we must resort to computer simulations to find out what it predicts. To apply it to the RRSP, we modeled the simulations after the experiments as closely as possible.

The simulations were run as follows. We always included both the control (nonreplacement) and the experimental (replacement) conditions. On each replication there was a study list of L items followed by a test list of $2L$ items, half new and half old. Presentation order for both the study phase and the test phase of each trial was always randomized. For the study phase the items were added to the memory vector in the prescribed fashion (Equation 2). That is, the dot product of the item with the memory vector was first computed, then the weighting factor c was simply the complement of this dot product. In some simulations we added information in the same way to the memory vector during the test phase; in other simulations we did not. Whether information was added to the memory vector during testing did not seem to affect the pattern of results we obtained.

In the experimental condition target items had to be replaced, and their replacements were simply more items of the same type. That is, all items of whatever type were simply N random samples from the feature distribution

Table 3
Trial 1 d' Values and the Hit-Rate Crossover on Trial 2 for Experiments 1-3

Experiment	FAR	HR	d'	Hit-Rate Crossover
1	.100	.765	1.996	.19
2	.155	.760	1.713	.17
3	.310	.599	0.744	.36

Note—FAR is mean false-alarm rate; HR is mean hit rate.

(i.e., a zero-centered normal distribution with $\sigma^2 = 1/N$). It was necessary to have a criterion to decide which items were not recognized on a given trial. In general, the criterion a was set about midway between zero and the mean of the old-item distribution, and to mimic the experimental data we sometimes raised the criterion over trials in the control condition. In general, these criterion manipulations seemed quite inconsequential. Although they certainly had a big effect on the number of targets, within the range we used they had little or no effect on the pattern of results.

In each replication there were two to five trials, depending on the simulation, and recognized items were replaced or not as the condition dictated. The new items were always different on each test trial. The value of N, the number of elements in the item vectors, ranged from 64 to 500, and the number of study items varied from 8 to 32, depending on the simulation. Although these lists were not as long as those used in the experiments, longer lists would have increased the running time for the simulations, which was already considerable (from 3 to 72 h per simulation on an IBM/XT with an 8087 coprocessor).

Finally, for each condition we initialized the memory vector to a set of random numbers before the first replication and left it alone from then on. That is, information was added during the study and test phases, but the memory vector was not reinitialized until we started the next condition (experimental or control). This means that the information in memory was cumulative over the 100-500 replications we ran per condition. Just as subjects remember items from test trials, so it is reasonable to assume that subjects remember items from prior lists. In fact, we have some unpublished data from a single-trial study-test procedure showing very clearly that subjects remember lists over sessions even when the sessions are spaced one or more days apart, so this cumulative procedure mimics in a small way what probably goes on naturally over periods of time spanning hours if not days.

The dependent variable was the dot-product value for each item on each trial, as this is the theoretically relevant measure. However, to identify the target items, it is necessary to assign "yes" and "no" responses to each item, so we determined this by whether the dot product was above or below the criterion for that trial. This enabled us to identify the targets and replace the nontargets in the study phase of the next trial in the experimental condition.

There were two main questions the simulation was designed to answer. The first question was whether learning occurred. There is little doubt that it did. Considering only the control condition, in all the simulations there was seldom a time when the mean dot-product values did not increase steadily with repetition. (The only exceptions were when the number of items in the list were about the same as the number of elements in the memory vector, and poor performance here would be expected.)

This is not enough to ensure that learning occurred, since we must consider the variances as well. Generally, the new-item variances increased very slightly over trials,

but the old-item variances *decreased* markedly over trials. Typically, the decrease in variance was about an order of magnitude (10:1) over say three to four trials, and this is a very large decrease. But then this is what a closed-loop system is intended to do: to "home in" on some desired value and achieve equilibrium. Thus, it seems safe to conclude that the open-loop system can learn.

The other question was whether the model generated a replacement effect. Sometimes there was a replacement effect, more often there was not, but on closer inspection of the results of the simulations this problem was replaced by a larger problem. The simulations clearly and consistently failed to yield a repetition effect. That is, in the experimental condition, targets and nontargets were of comparable strength. The results of the simulations, along with the relevant parameter values, are shown in Table 4 and, as can be seen, the mean strength of targets and nontargets was essentially the same.

We find, then, that the closed-loop version cannot handle one of the very striking results of the experiments, namely, that in the experimental group targets were better recognized than nontargets. As has been noted, not only was this repetition effect significant in each experiment, but there was only 1 subject in all three experiments who did not show a difference in this direction.

We turn now to a second modification of the matched-filter model, one based on probabilistic encoding, which does a better job of describing the main results of our experiments.

Probabilistic Encoding

The other modification of the matched-filter model that we consider is one in which each feature of an item is encoded with some fixed probability, p. This means that, on every trial, each feature of the item is or is not added to the memory vector. The dimensionality of the item vectors is not changed; rather, the encoding is probabilistic. This idea has been used by others (e.g., Estes, 1959;

Table 4
Frequencies, Means, and Standard Deviations of Targets and Nontargets for 16 Simulations (Sim) of the Closed-Loop Model

					Targets			Nontargets		
Sim	Rep	N	L	a	Freq	M	SD	Freq	M	SD
1	200	64	8	.3	1.1	.85	.44	6.9	.86	.50
2	200	64	16	.3	4.1	.77	.56	11.9	.74	.62
3	200	64	16	.3	3.8	.72	.59	12.2	.74	.63
4	100	200	30	.3	3.0	.84	.46	27.0	.82	.44
5	100	200	30	.5	6.0	.83	.41	24.0	.84	.41
6	100	200	30	.7	10.5	.84	.37	19.5	.84	.40
7	100	200	30	.5	5.9	.83	.41	24.1	.84	.41
8	100	200	30	.3	3.0	.84	.46	27.0	.82	.44
9	100	300	30	.3	1.2	.89	.33	28.8	.89	.35
10	100	400	30	.3	0.5	.96	.26	29.5	.91	.30
11	100	500	30	.3	0.3	1.06	.22	29.7	.93	.26
12	100	500	32	.3	0.4	.94	.24	31.6	.92	.28
13	300	500	32	.3	0.4	.94	.24	31.6	.92	.28
14	250	500	32	.5	1.8	.93	.26	30.2	.92	.26
15	250	500	32	.5	1.6	.92	.24	30.4	.93	.26
16	300	400	32	.45	1.9	.92	.28	30.1	.91	.30
Mean						.881	.374		.869	.402

Note—Rep is number of replications; L is list length.

Table 5
Hit-Rate Crossover on Trial 2 from a Simulation with $\alpha=.98$, $p=.4$, $N=200$, $L=15$, and a Yes–No Criterion at 0.3

Condition	Target	Nontarget
Control	.747	.987
Experimental	.940	.765

Hintzman, 1986), and it has some intuitive appeal. More important, it works. It allows learning to occur, and it seems to generate a replacement effect.

Unlike the closed-loop version, explicit expressions for the old-item mean, the new-item variance, and, consequently, the signal-to-noise ratio can be obtained. The derivations are presented in Appendix C. It can be seen that this is a more general version of the matched-filter model which reduces to the matched-filter model when $p=1$.

The rate of learning depends jointly on list length and the three parameters p, α, and N, but we do not analyze the model in any detail here. All we would like to do is present some sample data to show that the model can indeed generate a replacement effect. We have not done an extensive grid search to explore the parameter space; the results we present are illustrative, but the model is not yet complete.

The results from a simulation with $\alpha=.98$, $p=.4$, $N=200$, and a list length of 15 are shown in Table 5. The yes–no criterion was set at 0.3, order of presentation was randomized, and there were 100 replications. For both the control and the experimental group we assumed that the criterion was placed midway between the new- and the old-item mean, and the old-item variance was computed separately for targets and nontargets in each condition.

We did not use a parameter-estimation routine to obtain this crossover; it is the result of a fairly coarse grid-search with a few sample values of the criterion (.1 or .3), α (.95–.99), and p (.20–.80 in steps of about 0.5). (Remember, the criterion is the criterion in the simulation to segregate targets and nontargets on Trial 2, not the criterion used to determine the hit and false-alarm rates from the dot-product values in the simulation.)

Although the results of this simulation (and most others) generated a reasonable hit-rate crossover, quantitatively speaking they were not very close to the experimental data. We have not tried to improve the fit because the model is incomplete. In particular, it does not include any output interference. Information is added to the memory vector at the time of study but not at the time of test. From what we know of recognition memory, this is surely a major omission, and rectifying this omission is the next step to take in the further development of the model.

All we can claim, then, is that the model generates the right qualitative effect. In addition, it certainly demonstrates a repetition effect, and this was the downfall of the closed-loop model. In about 30 simulations it only failed to appear once, and generally the mean dot-product value for targets in the experimental condition was 20%–50% greater than the mean dot-product value for nontargets in the experimental condition. There was no

tendency for the magnitude of the hit-rate crossover to increase with task difficulty when difficulty was manipulated either by variations in α or by variations in p. This is admittedly a potential problem, but perhaps this problem could be solved by including output interference in the model.

Perhaps the most interesting result from the simulations was the context effect. There was no context effect at all when the proportion of targets was say 0.30 or above, but as the proportion decreased the context effect increased, and the lower the proportion the larger the effect became. That is, targets in the experimental condition had a higher dot-product value than targets in the control condition, and in the most extreme simulation the proportion of targets was about 0.025 and the context effect was almost 2:1.

Why does the model generate a context effect? After all, the same target items are repeated on Trial 2 in the replacement and nonreplacement conditions, so the dot-product value of the probe with these target items must be the same in both conditions. The answer probably lies in a consideration of the dot-product value of the probe with the nontargets. These must be sufficiently negative to offset the positive values resulting from the presence of the target item in the memory vector. Perhaps the intuitive argument about regression is correct after all; by replacing the nontargets, the targets benefit. It is not that the targets per se are any stronger in the experimental than in the control condition. Rather, the dot-product value is larger because the "masking" items are replaced by other items that are likely to be neutral rather than negative.

CONCLUSION

Our initial intuition turned out to be correct: there is a replacement effect and, using the RRSP, we were able to demonstrate a statistically significant hit-rate crossover in all three experiments. We also demonstrated a significant repetition effect but not a significant context effect. The matched-filter model of Anderson (1973) was shown not to be able to learn, and a closed-loop version could not generate the repetition effect. A second modification based on probabilistic encoding of item features was better in that it could generate both a hit-rate crossover and a repetition effect that was qualitatively consistent with the data.

We cannot claim complete success in modeling the data. We did not find a significant context effect in the experiments, and the probabilistic encoding model was not a good quantitative fit to the hit-rate crossover. Also, it is not clear whether or not the model can predict the increase in the magnitude of the hit-rate crossover with the increase in task difficulty that we found in these experiments. However, the model is incomplete and with the addition of output interference may fare better. Finally, the model provides some guidance as to how we might find a context effect. The smaller the proportion of targets, the

larger the context effect should be, and this provides an interesting challenge for further research.

REFERENCES

ANDERSON, J. A. (1973). A theory for the recognition of items from short memorized lists. *Psychological Review*, **80**, 417-438.

ANDERSON, J. A., SILVERSTEIN, J. W., RITZ, S. A., & JONES, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, **84**, 413-451.

ARBIB, M. A. (1964). *Brains, machines, and mathematics*. New York: McGraw-Hill.

AVANT, L. L., & BEVAN, W. (1968). Recognition of a stimulus class member after training with varied number of cases per class. *Journal of General Psychology*, **78**, 241-246.

CAVANAGH, J. P., & PARKMAN, J. (1971). A comparison of the study-recall and anticipation methods in steady-state paired-associate learning. *Psychonomic Science*, **22**, 361-363.

EICH, J. M. (1982). A composite holographic associative recall model. *Psychological Review*, **89**, 627-661.

EICH, J. M. (1985). Levels of processing, encoding specificity, elaboration, and CHARM. *Psychological Review*, **92**, 1-38.

ESTES, W. K. (1959). Component and pattern models with Markovian interpretations. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory* (pp. 9-52). Stanford, CA: Stanford University Press.

GLANZER, M., & BOWLES, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 21-31.

HEATH, R. A. (1983). An adaptive filter model for recognition memory. In J. A. Keats & R. A. Heath (Eds.), *Mathematics and behaviour: Conference Proceedings, 1982*. Newcastle, Australia: University of Newcastle.

HINTON, G. E., & ANDERSON, J. A. (Eds.). (1981). *Parallel models of associative memory*. Hillsdale, NJ: Erlbaum.

HINTZMAN, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, **93**, 411-428.

HUMPHREYS, M. S., PIKE, R., BAIN, J. D., & TEHAN, G. (in press). Global matching: A comparison of the SAM, Minerva, matrix, and TODAM models. *Journal of Mathematical Psychology*.

KAWAMOTO, A. H., & ANDERSON, J. A. (1985). A neural network model of multistable perception. *Acta Psychologica*, **59**, 35-65.

KINTSCH, W. (1965). The effects of repetition on the short-term memory function. *Psychonomic Science*, **2**, 149-150.

KNAPP, A. G., & ANDERSON, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 616-637.

KUČERA, H., & FRANCIS, W. N. (1967). *Computational analyses of present-day American English*. Providence, RI: Brown University Press.

LAMON, M. (1982). *The effects of context on recognition learning*. Unpublished master's thesis, University of Toronto.

LEWANDOWSKY, S., & HOCKLEY, W. E. (1987). Does CHARM need depth? Similarity and levels of processing effects in cued recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 443-455.

McCLELLAND, J. L., & RUMELHART, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, **114**, 159-188.

McDOWD, J. M., & MURDOCK, B. B., JR. (1986). Mathematical models of memory and the problem of stimulus variation: A comparison of MINERVA 2 and TODAM. *Acta Psychologica*, **62**, 177-188.

MURDOCK, B. B., JR. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, **89**, 609-626.

MURDOCK, B. B., JR. (1983). A distributed memory model for serial-order information. *Psychological Review*, **90**, 316-338.

MURDOCK, B. B., JR., & DUFTY, P. O. (1972). Strength theory and recognition memory. *Journal of Experimental Psychology*, **94**, 284-290.

NILSSON, L.-G. (1986). An interactionistic approach to models of distributed memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performance* (pp. 361-368). Amsterdam: Elsevier.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. (1968). Concreteness, imagery, and meaningfulness values for 125 nouns. *Journal of Experimental Psychology*, **76**(1, Pt. 2).

ROCK, I. (1957). The role of repetition in associative learning. *American Journal of Psychology*, **70**, 186-193.

SUTTON, R. S., & BARTO, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, **88**, 135-170.

WICKELGREN, W. A., & NORMAN, D. A. (1986). Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, **3**, 316-347.

WOOLDRIDGE, D. E. (1963). *The machinery of the brain*. New York: McGraw-Hill.

NOTES

1. We would like to thank W. K. Estes, one of the reviewers of an earlier version of this paper, for his comments on this problem.

2. Unfortunately not all of the raw data for the control condition of Experiment 1 were still available for this analysis, so in computing the value of the *t* test we used an estimate of the standard deviation whose ratio to that of the experimental group was twice as large as the comparable ratio in either Experiment 2 or Experiment 3.

3. We would like to thank M. J. Humphreys, one of the reviewers of an earlier version of this paper, for emphasizing the importance of ruling out this alternative explanation.

APPENDIX A

For a list of L items, each presented once, the mean of the old-item distribution $\mu_k(L)$, or simply μ_k, is given by:

$$\mu_k = E\left[\mathbf{f}_k \cdot \mathbf{M}\right] = E\left[\mathbf{f}_k \cdot \sum_{l=1}^{L} a_l \mathbf{f}_l\right]$$

$$= E\left[\mathbf{f}_k \cdot \left(a_k \mathbf{f}_k + \sum_{l \neq k}^{L} a_l \mathbf{f}_l\right)\right]$$

$$= E\left[\mathbf{f}_k \cdot a_k \mathbf{f}_k + \mathbf{f}_k \cdot \sum_{l \neq k}^{L} a_l \mathbf{f}_l\right]$$

$$= E\left[\mathbf{f}_k \cdot a_k \mathbf{f}_k\right] + E\left[\mathbf{f}_k \cdot \sum_{l \neq k}^{L} a_l \mathbf{f}_l\right]$$

$$= a_k E\left[\mathbf{f}_k \cdot \mathbf{f}_k\right] = N a_k E\left[Z^2\right]$$

$$= N a_k \sigma^2 = N a_k (P/N) = a_k P$$

for an item in serial position k, where σ^2 is the variance of the feature distribution. The probe item (from serial position k) is $\mathbf{f}_k$, and $\mathbf{M}$ is the memory vector. We use a_l as the serial-position constant, $l = 1, \ldots, L$, and it is a geometric function of α. Since $\mathbf{f}$ is a random vector, it can be considered as an ordered list of N random variables, each a random sample from the feature distribution. We use Z to denote the random variable. Since the items are independent, $\mathbf{f}_k \cdot \mathbf{f}_l = 0$, $l \neq k$. Thus, if $P = 1$ (P is the power of the vector) then $\mu_k = a_k$ and, from Equation 1, $a_k = \alpha^{L-k}$.

For R presentations of this list with a fixed order of presentation, the mean of the old-item distribution $\mu_o(R,L)$, or simply μ_o, is:

$$\mu_o = \frac{1}{L} \sum_{r=1}^{R} \sum_{l=1}^{L} \mu_l = \frac{1}{L} \sum_{r=1}^{R} \sum_{l=1}^{L} c_{r,l}$$

$$= \frac{1}{L} \frac{1 - \alpha^{RL}}{1 - \alpha^L} \sum_{l=1}^{L} \alpha^{L-l} = \frac{1}{L} \frac{1 - \alpha^{RL}}{1 - \alpha^L} \frac{1 - \alpha^L}{1 - \alpha}$$

$$= \frac{1}{L} \frac{1 - \alpha^{RL}}{1 - \alpha}.$$

We use $c_{r,l}$ rather than a_l to emphasize the fact that the serial-position constant varies with both repetition (r) and serial position (l). Assuming a fixed order, we can factor out the repetition component, which is simply the geometric sequence $1, \alpha^L, \alpha^{2L}, \ldots, \alpha^{(R-1)L}$, and we use the standard expression for the sum of the first R terms of a geometric progression. We then do likewise for the serial-position component, which is $1, \alpha, \alpha^2, \ldots, \alpha^{L-1}$.

For one presentation of the list, the new-item variance $\sigma_N^2(L)$, or simply σ_N^2, for the probe item $\mathbf{g}$ dotted with the memory vector $\mathbf{M}$ is:

$$\sigma_N^2 = \text{Var}\left[\mathbf{g} \cdot \mathbf{M}\right] = \text{Var}\left[\mathbf{g} \cdot \sum_{l=1}^{L} a_l \mathbf{f}_l\right]$$

$$= \sum_{l=1}^{L} a_l^2 \text{Var}\left[\mathbf{g} \cdot \mathbf{f}\right] = N \sum_{l=1}^{L} a_l^2 \text{Var}\left[ZW\right]$$

$$= N \sum_{l=1}^{L} a_l^2 \sigma^4 = N \sum_{l=1}^{L} a_l^2 (P^2/N^2),$$

where Z and W are identically distributed independent random variables. We use the fact that

$$\text{Var}\left[\sum_{l=1}^{L} c_l Z\right] = \text{Var}\left[c_1 Z + \ldots + c_L Z\right]$$

$$= \text{Var}\left[c_1 Z\right] + \ldots + \text{Var}\left[c_L Z\right]$$

$$= c_1^2 \text{Var}\left[Z\right] + \ldots + c_L^2 \text{Var}\left[Z\right]$$

$$= \sum_{l=1}^{L} c_l^2 \text{Var}\left[Z\right],$$

where the c_l's are constant terms. If $P = 1$, then, for a single presentation of the list, the new-item variance σ_N^2 is:

$$\sigma_N^2 = \frac{1}{N} \sum_{l=1}^{L} a_l^2.$$

For R presentations of the list, we need $\sigma_N^2(R,L)$, or simply σ_N^2, which is:

$$\sigma_N^2 = \text{Var}\left[\mathbf{g} \cdot \mathbf{M}\right] = \text{Var}\left[\mathbf{g} \cdot \sum_{r=1}^{R} \sum_{l=1}^{L} c_{r,l} \mathbf{f}_l\right].$$

Assuming a fixed order of presentation, we can write $c_{r,l}$ as the product of a_l and b_r, where a_l varies with serial position but not with repetition, and b_r varies with repetition but not with serial position. Consequently,

$$\text{Var}\left[\mathbf{g} \cdot \sum_{r=1}^{R} \sum_{l=1}^{L} c_{r,l} \mathbf{f}_l\right] = \text{Var}\left[\mathbf{g} \cdot \sum_{r=1}^{R} b_r \sum_{l=1}^{L} a_l \mathbf{f}_l\right]$$

$$= \left(\sum_{r=1}^{R} b_r\right)^2 \text{Var}\left[\mathbf{g} \cdot \sum_{l=1}^{L} a_l \mathbf{f}_l\right]$$

$$= \left(\sum_{r=1}^{R} b_r\right)^2 \text{Var}\left[\sum_{l=1}^{L} a_l \mathbf{f}_l \cdot \mathbf{g}\right]$$

$$= N\left(\sum_{r=1}^{R} b_r\right)^2 \sum_{l=1}^{L} a_l^2 \text{Var}\left[ZW\right]$$

$$= N\left(\frac{1 - \alpha^{RL}}{1 - \alpha^L}\right)^2 \frac{1 - \alpha^{2L}}{1 - \alpha^2} \sigma^4$$

$$= \frac{1}{N}\left(\frac{1 - \alpha^{RL}}{1 - \alpha^L}\right)^2 \frac{1 - \alpha^{2L}}{1 - \alpha^2}.$$

The main point here is that the repetition component (b_r) factors out differently from the serial-position constant (a_l). To see this, consider a specific example with three repetitions of a three-item list, where the list elements are the (identically distributed) random variables X, Y, and Z. This is illustrated in Table A1. If you take the double summation over R and L, the repetition

Table A1
Illustrations of Three Repetitions (R) of a List of Three Random Variables X, Y, and Z with Coefficients b_r and a_i

R	List		
	1	2	3
1	b_1a_1X	b_1a_2Y	b_1a_3Z
2	b_2a_1X	b_2a_2Y	b_2a_3Z
3	b_3a_1X	b_3a_2Y	b_3a_3Z

factor comes out as $(\Sigma b_r)^2$, whereas the serial-position factor comes out as Σa_i^2. Remember that the items (here, random variables) are assumed to be the same on each presentation, but they vary randomly from one serial position to another.

Since the signal-to-noise ratio (S/N) is by definition μ_o^2/σ_N^2, we have:

$$S/N = \frac{\{(1/L)[(1-\alpha^{RL})/(1-\alpha)]\}^2}{(1/N)[(1-\alpha^{RL})/(1-\alpha^L)]^2[(1-\alpha^{2L})/(1-\alpha^2)]}$$

$$= (N/L^2)[(1-\alpha^L)/(1-\alpha)]^2[(1-\alpha^2)/(1-\alpha^{2L})]$$

$$= (N/L^2)[(1+\alpha)/(1-\alpha)][(1-\alpha^L)/(1+\alpha^L)].$$

Consequently d', which is the square root of the S/N ratio, will vary with α, L, and N, but it will not vary with R, the number of presentations. Thus, performance as measured by d' will not improve with repetition, so learning will not occur.

APPENDIX B

Let $\phi(s)$ be the item strength distribution after a single presentation of the list. For the control condition $\mu_C(T,1)$, the mean strength of the target items after one presentation is given by:

$$\mu_C(T,1) = d' + \int_{-\infty}^{a} s\phi(s)ds,$$

where a is the location of the criterion on the strength (s) axis and d' is the strength resulting from the single presentation. The mean strength of nontargets (N) in the control condition $\mu_C(N,1)$ is given by:

$$\mu_C(N,1) = d' + \int_{a}^{\infty} s\phi(s)ds.$$

These integrals can be evaluated since for a unit normal distribution

$$\int_{a}^{b} s\phi(s)ds = \frac{\phi(a)-\phi(b)}{\Phi(b)-\Phi(a)}$$

(Murdock & Dufty, 1972). Thus we have at once that

$$\mu_C(T,1) = d' - \frac{\phi(a)}{\Phi(a)}$$

and

$$\mu_C(N,1) = d' + \frac{\phi(a)}{1-\Phi(a)}.$$

Verbally, the mean strength of items below the criterion (i.e., targets) is proportional to the ordinate of the normal curve at the criterion divided by the area below the criterion, whereas the mean strength of items above the criterion (i.e., nontargets) is proportional to the ordinate divided by the complement of the area.

Let $\Delta R_C(T)$ be the repetition effect for targets in the control condition (i.e., the increase in strength resulting from a second presentation) and $\Delta R_C(N)$ be the same for nontargets. In the

control condition, then, the mean strength for targets after the second presentation $\mu_C(T,2)$ is simply

$$\mu_C(T,2) = d' - \frac{\phi(a)}{\Phi(a)} + \Delta R_C(T)$$

and for nontargets

$$\mu_C(N,2) = d' + \frac{\phi(a)}{1-\Phi(a)} + \Delta R_C(N).$$

Then the difference, D_C, between nontargets and targets in the control condition is:

$$D_C = \mu_C(N,2) - \mu_C(T,2)$$

$$= \frac{\phi(a)}{1-\Phi(a)} + \frac{\phi(a)}{\Phi(a)} + \Delta R_C(N) - \Delta R_C(T).$$

In the experimental condition (E) for targets

$$\mu_E(T,2) = d' - \frac{\phi(a)}{\Phi(a)} + \Delta R_E(T),$$

but for nontargets

$$\mu_E(N,2) = d' + \frac{\phi(a)}{1-\Phi(a)},$$

which is the same as $\mu_C(N,1)$ because these items are being presented for the first time. Then D_E, the difference between nontargets and targets in the experimental condition, is given by:

$$D_E = \mu_E(N,2) - \mu_E(T,2) = \frac{\phi(a)}{1-\Phi(a)} + \frac{\phi(a)}{\Phi(a)} - \Delta R_E(T).$$

The hit-rate crossover (H) is simply

$$H = D_C - D_E = \Delta R_C(N) - \Delta R_C(T) + \Delta R_E(T).$$

Now the critical assumption of strength theory is that $\Delta R_C(T) = \Delta R_E(T)$, that is, that the benefits of repetition are item specific and so should be the same for targets in the experimental and control conditions. If this assumption were correct, then

$$H = \Delta R_C(N).$$

Thus, a simple version of strength theory with the assumptions given above predicts that the magnitude of H should be equal to the increment in strength due to repetition. Since the benefits of repetition will surely decrease as task difficulty increases, strength theory predicts that the magnitude of H should decrease as task difficulty increases.

APPENDIX C

Let $M_i = \alpha M_{i-1} + pf_i$, where pf_i denotes that each feature of the jth item is encoded (added to the memory vector) with some probability p. Then, as in Appendix A, after R presentations of a list of L items, the old-item mean $\mu_k(R,L)$, or simply μ_k, is:

$$\mu_k = E[\mathbf{f}_k \cdot \mathbf{M}] = E\left[\mathbf{f}_k \cdot \Sigma_{r=1}^{R} \Sigma_{l=1}^{L} c_{r,l}pf_l\right]$$

$$= E\left[\mathbf{f}_k \cdot \left(\Sigma_{r=1}^{R} c_{r,k}pf_k + \Sigma_{r=1}^{R} \Sigma_{l\neq k}^{L} c_{r,l}pf_l\right)\right]$$

$$= a_k E\left[\mathbf{f} \cdot \Sigma_{r=1}^{R} b_r pf\right].$$

We have skipped a number of steps that are spelled out in Appendix A. As before, the $l \neq k$ summation term drops out be-

cause items are independent, so $E[f_k \cdot f_l] = 0$, $l \neq k$. We can factor $c_{r,k}$ into two parts a_k and b_r, where a_k varies with serial position but not repetition and b_r varies with repetition but not serial position. Thus, $\Sigma c_{r,k} = a_k \Sigma b_r$, and we can drop the subscript on f because k is arbitrary. Now

$$a_k E\left[f \cdot \sum_{r=1}^{R} b_r p f\right] = a_k E\left[\sum_{r=1}^{R} b_r p f \cdot f\right]$$

$$= a_k E\left[\sum_{r=1}^{R} p b_r f \cdot f\right]$$

$$= a_k E\left[\sum_{r=1}^{R} p b_r\right] E\left[f \cdot f\right]$$

$$= a_k p \sum_{r=1}^{R} b_r E\left[f \cdot f\right]$$

$$= p a_k \sum_{r=1}^{R} b_r N\sigma^2.$$

That is, we can consider the probability p to operate on the serial-position constant rather than on the elements of the vector; that is, $b(pf) = (bp)f = (pb)f$ and $E[((pb)f] = E[pb]E[f] = pbE[f]$. If $P=1$, then we have

$$\mu_k = p a_k \sum_{r=1}^{R} b_r = p a_k \frac{1-\alpha^{RL}}{1-\alpha^L}.$$

This expression for μ_k gives us the mean dot-product value when the probe is an old item from serial position k. For the old-item mean, μ_O, we have

$$\mu_O = \frac{1}{L} \sum_{l=1}^{L} \mu_l = \frac{p}{L} \frac{1-\alpha^L}{1-\alpha} \frac{1-\alpha^{RL}}{1-\alpha^L} = \frac{p}{L} \frac{1-\alpha^{RL}}{1-\alpha}.$$

For the new-item variance, σ_N^2, we have

$$\sigma_N^2 = \text{Var}[g \cdot M] = \text{Var}\left[g \cdot \sum_{l=1}^{L} a_l \sum_{r=1}^{R} b_r p f_l\right]$$

$$= \sum_{l=1}^{L} a_l^2 \text{Var}\left[g \cdot \sum_{r=1}^{R} b_r p f\right]$$

$$= \sum_{l=1}^{L} a_l^2 \text{Var}\left[\sum_{r=1}^{R} b_r p f \cdot g\right]$$

$$= N \sum_{l=1}^{L} a_l^2 \text{Var}\left[\sum_{r=1}^{R} b_r p Z W\right]$$

$$= N \sum_{l=1}^{L} a_l^2 \text{Var}\left[\sum_{r=1}^{R} p b_r Z W\right].$$

All we have done here is to bring the serial-position constant a_l outside the bracket and switch from random vectors to random variables. Now $\Sigma p b_r Z W$ is an expression for the sum of all $2^R - 1$ nonzero outcomes (e.g., for $R=3$, these outcomes are $b_1 + b_2 + b_3$, $b_1 + b_2$, $b_1 + b_3$, $b_2 + b_3$, b_1, b_2, and b_3) with binomial probabilities p^3, $p^2 q$, and pq^2, respectively, where $q = 1 - p$. Now

Table C1
Numerical Example to Illustrate the S/N Ratio for $\alpha = .98$, $N = 200$, and $L = 15$

R	p			
	1.0	.80	.40	.05
1	13.23	10.59	5.29	0.66
3	13.23	12.16	8.65	1.72
5	13.23	12.50	9.78	2.42

Note—R is the number of presentations; p is the probability of encoding.

$$\text{Var}\left[\sum_{r=1}^{R} p b_r Z W\right]$$

$$= \left\{p \sum_{r=1}^{R} b_r^2 + 2p^2 \sum_{r=1}^{R} \sum_{r'>r}^{R} b_r b_{r'}\right\} \text{Var}[ZW]$$

$$= \left\{p \sum_{r=1}^{R} b_r^2 + p^2 \sum_{r=1}^{R} \sum_{r'\neq r}^{R} b_r b_{r'}\right\} \sigma^4$$

$$= \left\{p \sum_{r=1}^{R} b_r^2 + p^2\left[\left(\sum_{r=1}^{R} b_r\right)^2 - \sum_{r=1}^{R} b_r^2\right]\right\} \sigma^4$$

$$= \left\{p^2\left(\sum_{r=1}^{R} b_r\right)^2 + pq \sum_{r=1}^{R} b_r^2\right\} \sigma^4$$

where, as above, $q = 1 - p$. Consequently, we have

$$\sigma_N^2 = N \sum_{l=1}^{L} a_l^2 \left\{p^2\left(\sum_{r=1}^{R} b_r\right)^2 + pq \sum_{r=1}^{R} b_r^2\right\} \sigma^4$$

$$= \frac{1}{N} \sum_{l=1}^{L} a_l^2 \left\{p^2\left(\sum_{r=1}^{R} b_r\right)^2 + pq \sum_{r=1}^{R} b_r^2\right\}.$$

This is the new-item variance for a list of L items presented R times with probabilistic encoding.

The expression for the S/N ratio can be written as:

$$S/N = \frac{pN}{L^2} \frac{1+\alpha}{1-\alpha} \frac{(1-\alpha^{RL})(1-\alpha^L)}{p(1-\alpha^{RL})(1+\alpha^L) + q(1+\alpha^{RL})(1-\alpha^L)}$$

so the S/N ratio is a function of p, N, α, L, and R. A numerical example is shown in Table C1.

As can be seen, the absolute performance level drops but learning (the rate of change of the S/N ratio with respect to R) increases as p decreases. This probabilistic-encoding model is a more general version of the matched-filter model, and the expressions for μ_O, σ_N^2, and the S/N ratio given here reduce to the comparable expressions in Appendix A when $p=1$.