*Journal of Experimental Psychology*
1971, Vol. 91, No. 2, 220–226

# COMPARISON OF RECOGNITION AND RECALL IN A CONTINUOUS MEMORY TASK

GEOFFREY R. LOFTUS [1]

*Stanford University*

Differences between recall and recognition performance may be due at least in part to differences in the way information is stored. This possibility was explored in a paired-associate learning task by varying *S*'s knowledge at the time of study of how he would be tested on a particular stimulus–response pair. It was found that when *S* knew how he would be tested, his performance was better on recall but worse on recognition than when he did not know how he would be tested. These results were interpreted as support for the assertion that differences in storage processes partially account for recall-recognition performance differences. A model which postulates a distinction between short-term and long-term memory provided an excellent fit to the data and suggested possible storage strategies for recall and recognition.

Studies comparing recognition and recall have found performance differences between the two. When raw scores are considered, recognition is generally superior to recall (Freund, Brelsford, & Atkinson, 1969; MacDougall, 1904; Postman, 1950; Postman, Jenkins, & Postman, 1948. It can be assumed that this superiority is due, at least in part, to differences in processes taking place at the time of *retrieval*; in particular, the probability of guessing the correct response is higher in recognition than in recall. The major purpose of the present experiment is to investigate the extent to which recall-recognition differences in a paired-associate task may be attributed to differences in *S*'s method of *storing* information.

A convenient way to carry out such an investigation is to vary *S*'s knowledge at the time of study of how he will subsequently be tested on some item. If he lacks this knowledge, *S* will be forced to study all items in the same way, and recall-recognition differences on such items must be attributable to retrieval processes alone. Having the knowledge, on the other hand, enables *S* to store information differentially for recall and for recognition if he so desires. In the present experiment there are three conditions under which *S* may study an item: (*a*) He may study with the knowledge that he will be tested by recall; items in this condition are *recall* (Re) items. (*b*) He may study knowing that he will be tested by recognition; items in this condition are *recognition* (Ro) items. (*c*) Finally, *S* may study knowing that he will be tested either by recognition or recall but not knowing which; items in this condition tested by recognition are *recognition mixed* (RoM) items, and those tested by recall are *recall mixed* (ReM) items. Consider *S*'s performance when tested on the various types of items: If Ro does not differ from RoM and Re does not differ from ReM, the implication is that storage differences are not necessary to account for recall-recognition differences. To illustrate why this is so, consider a storage operator, **S**, which we use to represent various rehearsal schemes, mnemonic coding and, in general, any device used by *S* to store information for subsequent retrieval. The application of **S** to an input **I** represents information stored in memory, **S**(**I**). Further, a reretrieval operator, **R**, represents memory search, processing of stored information, and output of response at the time of test. Application of **R** to stored information produces output $p$, which is the probability of a correct response; i.e., $p = \mathbf{R}(\mathbf{S}(\mathbf{I}))$. For each of the three study conditions, there is a storage operator; these are designated $\mathbf{S}_{re}$, $\mathbf{S}_{ro}$, and $\mathbf{S}_{m}$ for recall, recognition, and mixed. Similarly, for the

[1] Requests for reprints should be sent to Geoffrey R. Loftus, now at New York University, Department of Psychology, 707 Broadway, New York, New York 10003.

two test conditions there are retrieval operators, $\mathbf{R}_{re}$ and $\mathbf{R}_{ro}$ for recall and recognition. Thus performance of the four types of output may be represented by:

$$\begin{aligned} \text{Ro—}p_{\text{Ro}} &= \mathbf{R}_{ro}(\mathbf{S}_{ro}(\mathbf{I})) \\ \text{RoM—}p_{\text{RoM}} &= \mathbf{R}_{ro}(\mathbf{S}_{m}(\mathbf{I})) \\ \text{Re}M\text{—}p_{\text{ReM}} &= \mathbf{R}_{re}(\mathbf{S}_{m}(\mathbf{I})) \\ \text{Re—}p_{\text{Re}} &= \mathbf{R}_{re}(\mathbf{S}_{re}(\mathbf{I})) \end{aligned}$$

If $p_{\text{Ro}} = p_{\text{RoM}}$ and if $p_{\text{Re}} = p_{\text{ReM}}$, then we may conclude that $\mathbf{S}_{ro}$ is the same as $\mathbf{S}_{m}$ and that $\mathbf{S}_{re}$ is the same as $\mathbf{S}_{m}$. Thus $\mathbf{S}_{re}$ and $\mathbf{S}_{ro}$ are identical, i.e., storage is the same for recall and recognition.

This paradigm was used by Freund et al. (1969) using paired-associate lists learned by the anticipation method. Their findings suggest that recall-recognition differences are due to retrieval alone; for both recall and recognition, varying *S*'s opportunity to store differentially did not lead to performance differences. Their design, however, may have been biased against finding storage differences for several reasons. First, their two-alternative forced-choice "recognition" test did not differ greatly from their nine-alternative forced-choice "recall" test. Second, and more important, *S*s were given all conditions within a single experimental session. The *S* would first study an item with the information that it would be tested by recall. This was followed by the study of a "recognition" item and then by a "test unknown" item. The cycle would then begin again with another "recall" item. To store differentially, *S* would have to change his method of study on a moment-to-moment basis. It is quite possible that this strategy of constant switching is too difficult to use.

The present study was similar in design to the Freund et al. (1969) experiment but it had two important changes. (*a*) A yes–no test was used for recognition. The response set consisted of the letters of the alphabet; thus "recall" was essentially from 26 alternatives. It was expected that these two tests would be more effective in demonstrating recall versus recognition differences than the 2- and 9-alternative forced-choice tests used by Freund et al. (1969). (*b*) Within an experimental session, *S* was in only one study condition: recall, recognition, or mixed. A single storage strategy could therefore be used throughout a session without eliminating the possibility of storing differentially as a function of condition.

## Method

*Subjects.*—The *S*s were eight female graduate students who received $2.00 for each session.

*Apparatus.*—The control functions of the experiment were performed by a computer program running on an on-line, modified, PDP-1d computer manufactured by Digital Equipment Corporation. The program operated on a time-sharing basis to drive eight KSR-33 teletypes which were situated in a single, windowless, soundproofed room in another building. Each *S* sat at a teletype equipped with a standard keyboard and a continuous roll of paper, masked in such a way that a horizontal strip about ½ in. wide was all that would be seen at a given time.

*Procedure.*—The stimuli were the digits 1–9, and the responses were the 26 letters of the alphabet. The *S*s served in one experimental session per day which took approximately 60 min. A within-*S*s design was used; each *S* was in each of the three experimental conditions five times and was in only one condition per session. The experimental conditions for each *S* were randomized over sessions, with the restrictions that she would be in all three conditions over a three-session block and would never be in the same condition for two sessions in a row. Each *S* had one initial practice session in the mixed condition.

A session started when *S* hit a code key on the teletype. The teletype would then print out what type of condition *S* was in for that day: recall, recognition, or mixed. A continuous task was employed: the nine stimuli were initially paired with randomly selected responses, and then the test-study trials which constituted the bulk of the session began. A stimulus to be tested was chosen randomly; in the mixed condition the type of test (recall or recognition) was determined randomly. The *S* had 30 sec. to recall the correct response and type in the appropriate letter in a recall test or, in a recognition test, had 30 sec. to type one of two special keys marked "yes" or "no." In recognition tests, the program randomly decided whether to present the correct response or a foil with the stimulus; if a foil was chosen, it was picked randomly from the 25 incorrect letters. If *S* was correct, feedback in the form of "+ + +" was printed out; if she was incorrect, "– – –" was printed out. The stimulus would then be re-paired with another response chosen randomly with the restriction that it could not be the one which had just been the correct response. The *S* was given 2 sec. to study the new pairing after which two carriage returns caused all typed material to disappear behind a mask, and *S*

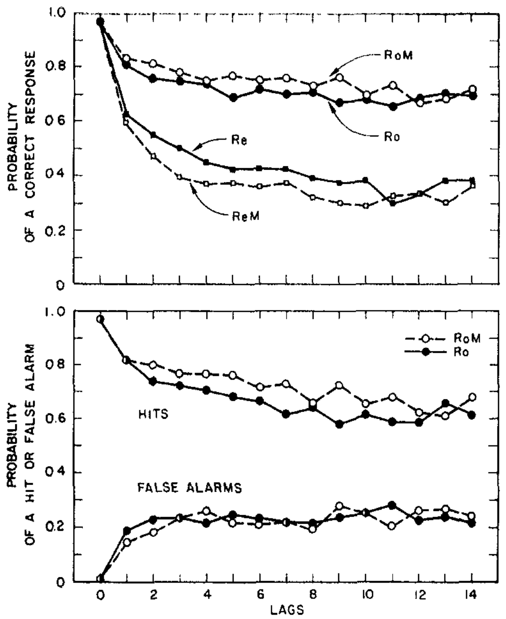


FIG. 1. Top panel—Probability correct as a function of lag for the Re, Ro, ReM, and RoM conditions. (Bottom panel—Hits and false alarms as functions of lag for the Ro and RoM conditions.)

was again tested on a stimulus chosen randomly from the set of nine digits. This procedure continued for 300 trials. At all times, *S*'s task was to remember the *last* response with which each stimulus had been paired. It should be emphasized that the program allowed each *S* to go at her own pace. After *S* responded to a test (or after 30 sec.) the program continued, and *S*'s feedback and the new pairing were immediately printed out.

In this type of continuous-task paradigm, the *lag* of an item being tested is defined as the number of test-study trials which have intervened since the item was last studied. The manner of selecting a stimulus to test results in a geometrically decreasing probability of being tested over lags. An item is tested immediately after being studied with probability 1/9, since the stimulus to be tested is chosen randomly out of 9, and, in general, the probability that the lag is equal to $i$ is $(8/9)^i(1/9)$, since the stimulus of interest is not chosen $i$ times (each time with probability 8/9) and then is chosen with probability 1/9.

## RESULTS

It has been found in previous studies using a continuous task (e.g., Atkinson, Brelsford, & Shiffrin, 1967) that a slight warm-up takes place at the beginning of a session. For this reason, the data from the first 25 trials of each session were excluded from the data analysis. In addition, the first session of data from each *S* was excluded, since this was regarded as a practice session during which *S*s were adapting to the task and the equipment.

The top panel of Fig. 1 presents the probability of a correct response as a function of lag for the Re, Ro, ReM, and RoM conditions. All curves decreased with increasing lag as expected, and overall performance was quite good. Even with 14 intervening items, performance was far above the chance levels of .04 for recall and .50 for recognition. For purposes of analysis, the recognition conditions are broken into "hits" (the probability of responding "yes" given that the correct response was presented) and "false alarms" (the probability of responding "yes" given that a foil was presented). These probabilities as functions of lag are presented in the bottom panel of Fig. 1. While the hit curves appear to steadily decrease, the false-alarm curves rise to an asymptote at about Lag 2 or 3 and then remain quite stable. For the Re and Ro curves, there are about 1,100 observations at Lag 0. Accordingly, for the ReM, RoM, and the hit and false-alarm curves corresponding to the Ro condition, there are about 550 observations at Lag 0, and for the hit and false-alarm curves corresponding to the RoM condition, there are about 275 observations at Lag 0.

Of crucial importance is the fact that differences exist between the Ro and RoM curves and between the Re and ReM curves. In order to test these differences, the response probability, pooled over all lags, was computed for each *S* for each curve. To compare any two curves, a $t$ test for matched pairs was performed. For Re versus ReM and for Ro versus RoM, both these tests were significant, $t(7) = 2.35$, $p \approx .05$ and $t(7) = 2.84$, $p < .05$. The two hit curves were found to be significantly different, $t(7) = 2.75$, $p < .05$, but the false-alarm curves did not differ significantly, $t(7) = .65$, $p > .20$. This indicates that the difference between the two recognition conditions is due primarily to the hit rates.

A problem in interpreting these results arises from the fact that the nature of the activity intervening between study and test of a given item differed in the various conditions. Only recognition tests intervened in the recognition condition and only recall tests intervened in the recall condition, while both types of tests intervened in the mixed condition. Conceivably, the differences between the Re and ReM items and between the Ro and RoM items could be explained by assuming that a recognition test produces more interference than a recall test. To investigate this possibility, the following analysis was performed: In the mixed condition, items were examined which had either all recall or all recognition tests intervening between the time the item was studied and the time it was tested. These items were further subdivided according to whether they themselves had been tested by recall or by recognition. For items tested by recognition, the response probabilities were .801 and .824, respectively, for all recall or all recognition tests intervening. For items tested by recall, the corresponding probabilities were .609 and .606. Neither the main effect of intervening activity nor the Intervening Activity × Type of Test interaction approached significance (both *F*s $< 1$). Therefore, the notion that intervening recognition tests generate more interference than intervening recall tests cannot be used to explain the present data.

The main results of this study provide support for the notion that storage differences exist between recognition and recall, contrary to the findings of Freund et al. (1969). In terms of the notation introduced above, the storage operator $\mathbf{S}_{re}$ is apparently more efficient than $\mathbf{S}_{m}$ which in turn is better than $\mathbf{S}_{ro}$. The latter is somewhat surprising since it might be expected that *S*s would be better at storing information when they are aware of the type of test to be employed. This hypothesis, however, is not supported by the fact that RoM is bettern than Ro.

Thus, in the present experiment, storage differences between recall and recognition appear quite clearly. It is now of interest to make a somewhat more detailed examination of processes occurring at the time of *retrieval*. It has been shown (Freund et al., 1969) that when storage factors are held constant, the relative superiority of recognition versus recall is highly dependent upon the type of procedure used to correct for the disparate guessing rates between the types of test. In the present experiment, application of two such procedures illustrates this dependency. The first, which has been used frequently, is to assume that if *S* does not know the correct answer, he guesses randomly among the response alternatives (Hilgard, 1951). In this case, let the probability of knowing the correct response be $p'$. The probability of *not* knowing the correct response and guessing correctly is $(1 - p')\left(\frac{1}{N}\right)$, where $N$ is the number of response alternatives (in this case, 2 for recognition and 26 for recall). The observed probability correct, $p$, is then

$$p = p' + \frac{(1 - p')}{N}.$$

In terms of our analysis, $p = \mathbf{R}(\mathbf{S}(\mathbf{I}))$, this correction for guessing is now regarded as follows: Whereas it was assumed above that $\mathbf{R}$ encompassed *both* recovery of information from memory and response production, we postulate a new operator, $\mathbf{R}'$, to be only the former process. In other words, $\mathbf{R}'(\mathbf{S}(\mathbf{I}))$ corresponds only to recovered information and can be identified with $p'$. Now let $\mathbf{G}'$ be the $p'$ to $p$ transformation described by the above equation. Thus,

$$\begin{aligned} p &= \mathbf{G}'(p') \\ &= \mathbf{G}'(\mathbf{R}'(\mathbf{S}(\mathbf{I}))). \end{aligned}$$

Since $\mathbf{G}'$ transforms $p'$ to $p$, $(\mathbf{G}')^{-1}$ may be applied to observed values of $p$ to obtain

$$\begin{aligned} p' &= (\mathbf{G}')^{-1}(p) \\ &= \mathbf{R}'(\mathbf{S}(\mathbf{I})). \end{aligned}$$

By comparing these $p'$ values ("corrected" probabilities) for the ReM and RoM con-

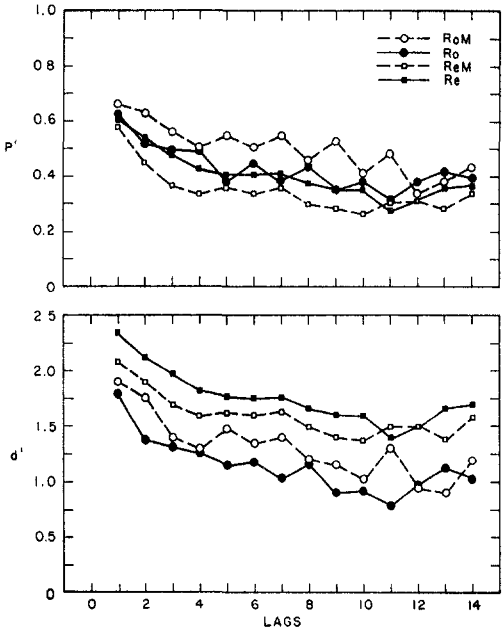


FIG. 2. Values of $p'$ (top panel) and $d'$ (bottom panel) for the Re, Ro, ReM, and RoM conditions.

ditions, we may directly compare our operators, $\mathbf{R}'_{re}$ and $\mathbf{R}'_{ro}$. Since the storage operators, $\mathbf{S}_m$ are the same, in these conditions, differences in $p'$ must be attributable only to differences in $\mathbf{R}'$. The top panel of Fig. 2 shows $p'$ as a function of lag for the Re, Ro, ReM, and RoM conditions. Performance on RoM is superior to that on ReM, $t(7) = 13.0$, $p < .01$. In terms of the above analysis, then, $\mathbf{R}'_{ro}$ leads to better performance than does $\mathbf{R}'_{re}$. The $p'$ curve for Ro was also above that for Re, although not significantly so, $t(7) = 1.16$, $p > .10$.

The second method for comparing recognition and recall is suggested by the theory of signal detectability (TSD) which has been successfully applied to recognition memory (Kintsch, 1968; Murdock, 1965). It is postulated that the strength of information about items in memory may be represented by a single value, $d'$. Given the probability correct for an $N$-alternative forced-choice test, or the hit and false-alarm rates for a yes–no test, it is possible to find the corresponding value of $d'$ (Elliott, 1964). Again, in terms of our operator notation, let $d' = \mathbf{R}''(\mathbf{S}(\mathbf{I}))$. Then,

$$\begin{aligned} p &= \mathbf{G}''(d') \\ &= \mathbf{G}''(\mathbf{R}''(\mathbf{S}(\mathbf{I}))), \end{aligned}$$

where $\mathbf{G}''$ is the transformation of $d'$ to response probability as specified in Elliott (1964). Again, $(\mathbf{G}'')^{-1}$ is explicitly defined and

$$\begin{aligned} d' &= (\mathbf{G}'')^{-1}(p) \\ &= \mathbf{R}''(\mathbf{S}(\mathbf{I})). \end{aligned}$$

The bottom panel of Fig. 2 presents $d'$ as a function of lag predicted from the data for the Re, Ro, ReM, and RoM conditions. The results are as follows: Re is now superior to Ro, $t(7) = 4.52$, $p < .01$, and ReM is superior to RoM, $t(7) = 4.24$, $p < .01$. Again examining the mixed conditions where the storage operators are the same $\mathbf{R}''_{re}$ leads to better performance than $\mathbf{R}''_{ro}$ which is a reversal of the findings which obtained when $(\mathbf{G}')^{-1}$, the $p'$ transformation, was applied to the data.

This reversal was also obtained by Freund et al. (1969). They applied the $p'$ and $d'$ corrections to their data and found that the former transformation showed recognition to be superior to recall, whereas the latter transformation showed the opposite to be true. Previous studies comparing recall and recognition have, in many cases, used the $p'$ correction for guessing and have concluded that recognition is superior to recall (Postman, 1950; Postman et al. 1948). Studies comparing recall and recognition using a TSD analysis have generally assumed that performance based on information acquired under identical conditions should lead to the same value of $d'$ independent of the type of test. As suggested above, however, it is not meaningful in the present experimental paradigm to make unconditional conclusions regarding the relative superiority of recall versus recognition. We have examined the view that some guessing correction, $\mathbf{G}$, defined for recall and recognition operates on retrieved information to produce a response. In this framework, application of $\mathbf{G}^{-1}$ to the data should produce "corrected" measures which are comparable for the two types of tests. To reiterate the conclusions of Freund et al., however, the choice of **G** and the resulting comparison of recall and recognition must rest on specific assumptions regarding the nature of the retrieval process.

## Discussion

We now turn to an analysis of the data in terms of a theory of memory which has been proposed by Atkinson and Shiffrin (1968). The model to be used in the present article represents an application of this theory to recognition memory and is described in detail by Freund et al. (1969) and Atkinson and Wickens (1971). Basically, the model postulates two memory states: a short-term store (STS) and a long-term store (LTS). Incoming information enters STS where it decays within a short period of time unless it is entered into and is maintained in a rehearsal buffer. Items are assumed to enter this buffer with probability $\alpha$. The buffer is assumed to have a fixed capacity of $r$ items; *S* may or may not enter a new item into the buffer, but if he does so, a randomly selected item currently in the buffer is knocked out and rapidly decays from STS. Information about an item is transferred to LTS from STS in two ways: (*a*) An amount of information $\theta'$ is initially transferred by virtue of the fact that the item entered STS, and (*b*) if the item enters the rehearsal buffer, additional information is transferred at a constant rate $\theta$ for each trial that the item resides in the buffer. After the item leaves the buffer, information about it in LTS decreases exponentially at a rate $1 - \tau$ per trial. For purposes of simplicity, $\theta$ is set equal to $\theta'$; thus the amount of available information at Lag $i$ about an item which had been in the buffer for $j$ trials ($j \leq i$) is equal to $(j + 1)\theta\tau^{i-j}$. If an item is in the rehearsal buffer when it is tested, a correct response is made with probability one. If it is not in the buffer, a response is made based on information retrieved from LTS. Here, the probability of a correct response is given by a TSD analysis where $d' = (j + 1)\theta\tau^{i-j}$. For yes–no tests, *S* has a response bias for responding "yes." This bias (or criterion), $c$, is another parameter of the model. For recall, the model thus has four parameters to be estimated: $\alpha$, $r$, $\theta$, and $\tau$. For recognition an additional parameter, $c$, is estimated.

The theory makes a distinction between structural features of memory (e.g., STS, LTS) and control processes used by *S* to deal with specific tasks. Different strategies (control processes) may lead to widely varying values of $\alpha$, $r$, and $\theta$. In the present experiment, storage differences may be thought of in terms of these control processes (Atkinson & Wickens, 1971); a different buffer size, for example, might be used by *S* depending on whether he knows he will be tested by recall or by recognition.

TABLE 1

Parameter Values and Chi-Squares for Three Conditions

| Parameter | Recognition | Mixed | Recall |
|---|---|---|---|
| $\alpha$ | .79 | .73 | .53 |
| $r$ | 1 | 2 | 3 |
| $\theta$ | .79 | .52 | .30 |
| | .95 | .97 | .99 |
| $c$ | .71 | .62 | — |
| $\chi^2$ | 22.3 | 29.3 | 11.3 |
| *df* | 23 | 37 | 10 |

Note.—A dash indicates that the parameter is not needed for the recall task.

A fit of the model to the observed data was made using a minimum $\chi^2$ procedure (Atkinson, Bower, & Crothers, 1964). Three separate parameter estimates were made: one for each of the three study conditions. For the recognition condition, the fit was made over the hit and false-alarm curves derived from the Ro condition, and for the recall condition, the fit was made to the Re curve. For the mixed condition, the parameters were estimated for the recall and the hit and false-alarm data simultaneously. The parameter estimates along with the corresponding $\chi^2$ values and degrees of freedom are shown in Table 1. In terms of the model, the parameter values suggest that *S*s used quite different control processes depending on the type of test they were anticipating. For the Ro condition, each item has a good deal of information about it transferred to LTS during one trial (the buffer size is 1 and $\alpha$ and $\theta$ are fairly high). For Re, on the other hand, there seems to be more emphasis on trying to maintain items in STS without a great deal of effort to store them in LTS (the buffer size is 3 and $\theta$ is low). For the mixed condition, as might be expected, an intermediate strategy is indicated (the buffer size is 2, and $\alpha$ and $\theta$ have intermediate values). The *S*s' verbal reports indicated that they were, in general, using the strategies suggested by the parameter values for the various conditions.

The different control processes for recall and recognition outlined above are logically consistent with the type of information needed for a recall as opposed to a recognition test. For recognition, minimal information about a response is often sufficient to generate a correct response. For example, if *S* can retrieve simply the fact that "the answer rhymes with A," this is enough to be correct if he is presented with "Q" and asked to respond yes or no. A good strategy would thus be to generate as much information as possible about each item and allow the information to decay away since it is still useful in a degraded form. For recall, on the other hand, such degraded information is not as useful, and there would be more reason to try to maintain complete information about as many items as possible in STS where it can be retrieved perfectly.

The strategy for recognition is similar to one suggested by the same model in a previous study. Freund et al. (1969) used a continuous-task paradigm analogous to the mixed condition of the present experiment. There were four types of tests: yes–no, 2, 4, and 26 forced choice, and *S* never knew at the time of study how an item would be tested. The yes–no, 2, and 4 forced-choice tests might well be regarded as recognition tests; thus a reasonable strategy, in terms of the above analysis, would be to store for recognition. When the model described above was applied to the data, the best fit was obtained with a buffer size of 1, $\alpha$ of .75, $\theta$ of .86, and $\tau$ of .95: These parameter values are remarkably similar to those obtained in the Ro condition of the present study.

There are several conclusions to be drawn from the findings of the present experiment. First, storage differences between recognition and recall have been found contrary to the findings of Freund et al. (1969). As indicated above, the major reason for this is probably that the methodology of the present study is better attuned to the separation of storage strategies. Second, analysis of the data in terms of the memory model discussed above has been highly successful. The model has provided an excellent fit to the observed data, and the obtained parameter values are interpretable in terms of possible storage strategies used by *S* in each of the three study conditions.

## REFERENCES

ATKINSON, R. C., BOWER, G. H., & CROTHERS, E. J. *An introduction to mathematical learning theory.* New York: John Wiley, 1964.

ATKINSON, R. C., BRELSFORD, J. W., JR., & SHIFFRIN, R. M. Multi-process models for memory with applications to a continuous presentation task. *Journal of Mathematical Psychology*, 1967, **4**, 277–300.

ATKINSON, R. C., & SHIFFRIN, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation.* Vol. 2. New York: Academic Press, 1968.

ATKINSON, R. C., & WICKENS, T. D. Human memory and the concept of reinforcement. In R. Glaser (Ed.), *The nature of reinforcement.* Columbus, O.: Merrill Publishing Co., in press.

ELLIOTT, P. B. Tables of *d'*. In J. A. Swets (Ed.), *Signal detection and recognition by human observers.* New York: John Wiley, 1964.

FREUND, R. D., BRELSFORD, J. W., JR., & ATKINSON, R. C. Recognition vs. recall: Storage or retrieval differences? *Quarterly Journal of Experimental Psychology*, 1969, **21**, 214–224.

FREUND, R. D., LOFTUS, G. R., & ATKINSON, R. C. Applications of multi-process models for memory to continuous recognition tasks. *Journal of Mathematical Psychology*, 1969, **6**, 576–594.

HILGARD. E. R. Methods and procedure in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: John Wiley, 1951.

KINTSCH, W. An experimental analysis of single stimulus tests and multiple-choice tests of recognition memory. *Journal of Experimental Psychology*, 1968, **76**, 1–6.

MACDOUGALL, R. Recognition and recall. *Journal of Philosophy*, 1904. **1**, 299–333.

MURDOCK, B. B. Signal detection theory and short-term memory. *Journal of Experimental Psychology*, 1965, **70**, 443–447.

POSTMAN, L. Choice behavior and the process of recognition. *American Journal of Psychology*, 1950, **63**, 576–583.

POSTMAN, L., JENKINS, W. O., & POSTMAN, D. L. An experimental comparison of active recognition and recall. *American Journal of Psychology*, 1948, **61**, 511–519.

(Received March 22, 1971)