

APPLICATIONS OF MULTIPROCESS MODELS FOR MEMORY TO
CONTINUOUS RECOGNITION TASKS

by

R. D. Freund, G. R. Loftus, and R. C. Atkinson

TECHNICAL REPORT NO. 138

December 18, 1968

PSYCHOLOGY SERIES

Reproduction in Whole or in Part is Permitted for
any Purpose of the United States Government

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

STATE OF TEXAS, COUNTY OF DALLAS

USDA, N. 101.00, 101.00, 101.00

and

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

USDA, N. 101.00, 101.00, 101.00

Abstract

Several multiprocess models for memory and learning are applied to the results of an experiment comparing performance on four types of recognition tests. The task involved a continuous sequence of trials, each trial consisting of a test on one of the stimuli followed by a study on that same stimulus paired with a new response. One of four types of tests was presented on each trial, the choice of test being made randomly. The four types of tests employed were a two, four, and twenty-six alternative forced-choice test, and a yes-no test. During the study period the subject had no way of knowing which mode of test would be given, and thus could not engage in differential storage processes. The basic dependent variable was the probability of a correct response as a function of the number of trials intervening between study and test on a given stimulus-response pair (called the "lag"). The lag curves for the forced-choice tests and the hit curve for the yes-no test decreased monotonically as the lag increased, while the false alarm curve of the yes-no test increased as the lag increased. A model which postulates a distinction between short-term and long-term memory stores was applied successfully to these data. The model assumes that information in short-term store is perfectly retrievable and utilizes an analysis derived from Signal Detectability Theory to describe long-term processes.

[Faint, illegible text, likely bleed-through from the reverse side of the page]

Applications of Multiprocess Models for Memory
to Continuous Recognition Tasks*

R. D. Freund, G. R. Loftus, and R. C. Atkinson

Stanford University

Stanford, California 94305

Atkinson and Shiffrin (1965, 1968a) have formulated a model for human memory that has been applied successfully to a wide range of tasks and experimental variables: list length, presentation rate, reaction times, and confidence ratings using discrete-trial memory tasks; mode of rehearsal, size of stimulus set, number of prior study trials, and temporal judgments in tasks involving a continuous sequence of study and test items. These previous studies employed measures of recall performance, either in free-verbal recall or paired-associate paradigms. The present study was designed to test the model in situations involving various types of recognition tests.

Differences between recognition and recall performance may be attributed to differences in storage processes, retrieval processes, or to some combination of both (Atkinson and Shiffrin, 1968b). The task employed in the experiment to be described here involves a modification of the typical paired-associate procedure which makes it possible to isolate the effects of retrieval processes in recognition tasks. This is accomplished by employing several types of recognition tests, but withholding information about the type of test to be given at the time

*This research was supported by Grant NGR-05-020-036 from the National Aeronautics and Space Administration.

the subject is studying a particular stimulus-response pair. Thus, there is no opportunity for storage to take place in one way in anticipation of one mode of test, and in some other way in anticipation of another test mode. There can be only one storage strategy, and therefore performance differences obtained among the different tests necessarily imply retrieval differences.

The task used was similar to the continuous paired-associate memory procedure described in Atkinson, Brelsford, and Shiffrin (1967). The subject is required to keep track of the randomly changing response members of eight different stimuli. Every trial is composed of a test period and study period on the same stimulus. During the test phase a stimulus is randomly selected from among the set of eight stimuli and one of four types of tests is presented, the choice of test being made randomly. On all tests the subject's task is to remember the response last associated with that stimulus. The tests employed were: (1) two-alternative forced-choice, where the subject must select the correct response from two alternatives; (2) four-alternative forced-choice, where the correct response must be selected from among four alternatives; (3) twenty-six alternative forced-choice, where the subject must select the correct response from among all possible alternatives; and (4) yes-no test, where one response is presented with the stimulus and the subject must indicate whether it is correct or incorrect. The distractors in the forced-choice tests were chosen randomly at the time of test. Similarly, the correct response in the yes-no test was presented on a random schedule; if an incorrect response was scheduled for presentation on a yes-no test, it was chosen randomly.

Following the test, the study phase of the trial occurs. During this phase the stimulus used in the test phase is re-paired with a new response for study. Crucial to the design is that the subject, at the time of study, does not know the mode of test to be employed the next time this stimulus is presented. This enables us to restrict our attention to processes occurring during the test phase--i.e., retrieval processes--in attempting to understand performance differences which occur among the four types of tests. The major performance measure obtained was the proportion of correct responses as a function of the "lag" at which an item was tested. The lag is defined as the number of trials intervening between the study and eventual test of a given stimulus-response pair. The choice of a given stimulus for test, as well as the mode of test to be employed, was determined on a random basis.

It should be clear that the subject is always trying to remember exactly eight stimulus-response pairs at any given time. Each time a stimulus is tested it is immediately re-paired with a new response, keeping the size of the to-be-remembered stimulus set always fixed at eight. Of course, in order to start an experimental session, an initial series of trials must be given with the test phase omitted. The stimuli presented during study trials are the ones used throughout the rest of the experimental session.

MODEL

The general model and the mathematical derivations of the lag curves for the continuous task are described in Atkinson, Brelsford, and Shiffrin (1967). Two specific versions of the general model will be presented here to deal with the recognition tasks. As indicated

above, the tasks we are concerned with are specifically ones which focus upon retrieval processes, rather than storage processes. Accordingly, the versions of the model to be presented are distinguished by specific assumptions about the retrieval of information from memory. This in no way implies that differential storage mechanisms could not be invoked in other tasks, nor that the model could not be amended to handle such differential storage processes. Rather, we have purposely limited our experimental and theoretical examination to retrieval processes, in the hopes of testing some of the more reasonable assumptions about recognition performance.

The general model postulates three memory states: a very short-lived state called the sensory register; a temporary memory state called the short-term store (STS); and a more permanent long-term store (LTS). Define an "item" as that amount of information which is sufficient to generate a correct response when the stimulus is presented for test. Items are postulated to enter and leave the various memory states at different times. Specifically, it is assumed that every item enters the sensory register and is immediately transferred to STS. Thus the sensory register plays no part in the application of the models to be discussed in this paper.

The short-term store is the subject's working memory, receiving information from the sensory register and from LTS, as well as directing the various control processes such as storage and retrieval. Information entered into STS will spontaneously decay within about 30 seconds, unless maintained by rehearsal processes activated by the subject. Such activity is represented by a rehearsal "buffer" having a limited and

constant capacity for homogeneous items. Specifically, r is the number of such items that can be maintained simultaneously in the buffer. Once filled, each new item which enters causes one of the items currently in the buffer to be lost. If the stimulus member of an item is already in the buffer, then the item is assumed to enter the buffer with probability one. Thus, it may be said to replace itself in the buffer. Alternatively, the stimulus member may not be in the buffer; then the item enters the buffer with probability α , knocking out some other item currently in the buffer. A random knock-out process for the present type of experiment has proved adequate (Atkinson, Brelsford, and Shiffrin, 1967) and shall be assumed here: the item to be knocked out is chosen independently of its duration in the buffer. It is assumed that a correct response is given with probability one if an item is in the buffer at the time of test.

Information about an item is assumed to accrue in LTS during the period that the item resides in the buffer. In general, this information represents the subject's encoded, storable representation of the stimulus-response pair. We assume that information is transferred to LTS at a constant rate θ during the period the item resides in the buffer. Furthermore, regardless of whether or not the item enters the buffer, each item transfers some fixed amount of information θ' to LTS by virtue of its entrance into and eventual decay from STS. That is, an item which does not enter the buffer is assumed to transfer information to LTS at a rate θ' during its decay from STS. Similarly, any item knocked out of the buffer will transfer information to LTS at this same rate during its decay period. Thus, the amount of information accrued in LTS for an

item which resided in the buffer for exactly j trials is: $j\theta + \theta'$. For simplicity we shall assume $\theta = \theta'$ in the present version of the model. Each trial following the trial on which an item is knocked out of the buffer causes the retrievable information stored about that item in LTS to decrease by a constant proportion τ . In summary, then, if i trials intervened between study and test on an item that resided in the buffer for j trials ($j \leq i$), then the amount of retrievable information in LTS for that item at the time of test is $(j+1)\theta\tau^{i-j}$.

At the time of test, the subject always gives the correct response if the item is in the short-term store. If the item is not in STS, the subject searches LTS and responds on the basis of the information about the item stored there. This LTS search is called the retrieval process. In general, the probability of a correct retrieval of an item from LTS is a monotonic increasing function of the amount of information about the item which is stored there. When no information is stored, the probability of a correct response should be at the guessing level. We define p_{ij} as the probability of a correct response from LTS of an item tested at a lag of i , which resided in the buffer for exactly j trials. Two characterizations of the retrieval function for the recognition tasks shall be presented. Since retrieval from STS is assumed to be perfect, any attempt to account for performance differences on the basis of retrieval processes will involve assumptions about LTS retrieval.

Model I

This model is identical to the one first postulated by Atkinson, Brelsford, and Shiffrin (1967), where a correction for guessing is combined with an exponential function relating retrieval to information in

LTS. Specifically, for the forced-choice conditions,

$$\rho_{ij}^{(N)} = 1 - \left(1 - \frac{1}{N}\right) \exp[-(j+1)\theta\tau^{i-j}] \quad (1)$$

where the superscript refers to the number of alternatives in the forced-choice test ($N = 2, 4, \text{ or } 26$). For the yes-no test, the analysis is divided into the probability of a "hit" (responding "yes" when the response displayed is in fact correct) and of a "false alarm" (responding "yes" when the response displayed is incorrect). Equation 1 may be generalized to include the hit and false alarm data as follows: assume some bias, γ , for responding "yes." Then the probability of a hit, given the item is not in the buffer, is identical to the probability of a correct retrieval from LTS in the forced-choice conditions, except that the guessing parameter, $1/N$, is replaced by γ . The expression then becomes

$$\rho_{ij}^{(\text{Hit})} = 1 - (1-\gamma) \exp[-(j+1)\theta\tau^{i-j}] \quad (2)$$

For false alarms, the subject must not retrieve the correct response, and then guess "yes" when the incorrect alternative is displayed. The expression is

$$\rho_{ij}^{(\text{FA})} = \gamma \exp[-(j+1)\theta\tau^{i-j}] \quad (3)$$

Model II

The application of the Theory of Signal Detectability (TSD) to recognition memory tasks has been considered in several reports (e.g., Wickelgren and Norman, 1966; Bernbach, 1967; Kintsch, 1967). Donaldson and Murdock (1968) recently discussed the rationale for applying TSD analysis to a continuous short-term recognition task. The present

treatment focuses the analysis on long-term processes. Specifically, we postulate that d' (the subject's sensitivity parameter) is directly related to the amount of information stored about an item in LTS. The simplest assumption is to let d' represent the current amount of information about an item in LTS. As described above, the amount of information stored in LTS for an item depends upon the number of trials between its study and test, and also upon the number of trials the item resided in the buffer. Therefore, we define

$$d'_{ij} = (j+1)\theta\tau^{i-j}$$

as the measure of the subject's sensitivity for an item which was tested at lag i and resided in the buffer j trials. Thus, θ in this formulation becomes a parameter of proportionality relating the amount of information in LTS to the subject's sensitivity in recognizing the correct response. Tables by Elliot (see Swets, 1964, pp. 682-683) allow conversion of observed response proportions to d'_{ij} as a function of the number of response alternatives in our two and four forced-choice conditions; for the 26-alternative forced-choice condition, the linear approximation given by Elliot (see Swets, 1964, p. 680) can be used. For the yes-no condition, the hit and false-alarm probabilities are

$$p_{ij}^{(\text{Hit})} = \int_c^\infty N(d'_{ij}, 1) \quad (4)$$

$$p_{ij}^{(\text{FA})} = \int_c^\infty N(0, 1) \quad (5)$$

where $N(i, 1)$ refers to a normal density function with mean i and unit variance.

Derivation of Lag Curves

Having specified the general model and the two retrieval functions to be evaluated, it remains to derive the probability of a correct response at any given lag i . Correct responses arise from STS and from LTS. We define β_j as the probability that an item resides in the buffer for exactly j trials, given that it is tested at a lag greater than j . Then the probability that an item is in the buffer, when tested at lag i , is $1 - (\beta_0 + \beta_1 + \dots + \beta_i)$ where β_0 is simply the probability that the item does not enter the buffer. One variation in the procedure of the present experiment from that reported in Atkinson, Brelsford, and Shiffrin (1967) should be mentioned here. It was decided to eliminate all tests at lag 0, since previous studies had demonstrated perfect performance with no intervening items between study and test. The omission of lag 0 tests made the derivations of β_j somewhat complex; they are presented in the Appendix.

For the N -alternative forced-choice conditions, the probability of a correct response from LTS of an item tested at lag i is

$$\sum_{j=0}^i \beta_j \rho_{ij}^{(N)}$$

In this expression $\rho_{ij}^{(N)}$ represents the probability of a correct response for the appropriate type of test ($N = 2, 4, \text{ or } 26$). Therefore, the overall probability of a correct response at lag i in the N -alternative forced-choice condition is

$$\Pr(C_i^{(N)}) = [1 - \sum_{j=0}^i \beta_j] + [\sum_{j=0}^i \beta_j \rho_{ij}^{(n)}] \quad (6)$$

For yes-no tests, the second bracketed term is replaced as follows: the probability of a hit is

$$\text{Pr}(\text{Hit at lag } i) = [1 - \sum_{j=0}^i \beta_j] + [\sum_{j=0}^i \beta_j \rho_{ij}^{(\text{Hit})}] \quad (7)$$

and the probability of a false alarm is

$$\text{Pr}(\text{FA at lag } i) = \sum_{j=0}^i \beta_j \rho_{ij}^{(\text{FA})} \quad (8)$$

METHOD

Subjects. The subjects were ten Stanford students who received two dollars per session. The group consisted of seven females and three males, all of whom participated in at least fourteen experimental sessions. Some of the subjects had previous experience in verbal learning tasks.

Apparatus. The experiment was conducted in the Computer-Based Learning Laboratory at Stanford University. The control functions were performed by a computer program running on a modified PDP-1 computer manufactured by the Digital Equipment Corporation, all under the control of a time-sharing system. The subject sat in an individual soundproof 7 x 9 ft. booth facing a cathode-ray tube (CRT), which displayed the stimuli. A standard typewriter keyboard, located directly beneath the lower edge of the CRT, was used for responding.

Stimuli and responses. The stimuli were two-digit numbers with certain high association ones removed (10, 20, 30, ..., 90; 11, 22, 33, ..., 99; 50, 98). For each subject and session a set of eight such

stimuli were selected at random, and this set was used throughout the session for that subject. The responses were the letters of the alphabet.

Procedure. Each experimental session began with a sequence of eight study trials, one for each stimulus to be used in the session. On a study trial the word study appeared on the upper face of the CRT. Beneath the word study one of the stimuli appeared along with a randomly-selected response. Subjects were instructed to try to remember the stimulus-response pairing. Once there had been an initial study trial for each stimulus, the session proper began.

Each of the subsequent two hundred trials of a session consisted of a test phase and a study phase, as follows: (1) The word test appeared on the upper face of the CRT. One of the eight stimuli was randomly chosen for test, with the restriction that it was not the stimulus chosen for test on the immediately preceding trial. This stimulus was displayed beneath the word test on the CRT. The type of test was determined randomly at this time. (a) Yes-No: a single response was displayed on the CRT to the right of the stimulus. A random decision was made as to whether this response was to be the correct response or a randomly-chosen distractor. The subject was instructed to respond "yes" if he thought the response displayed was the one which had last been associated with the stimulus, to respond "no" if he thought it was incorrect, and to guess if necessary. Two special keys, labeled "yes" and "no" were used for these responses. (b) Two-alternative forced-choice (FC-2): the two response alternatives were displayed to the right and slightly below the stimulus on the CRT. These responses were located side by side, separated by about $1\frac{1}{2}$ in. on the screen. The position of the correct response was

randomly determined, as was the choice of the incorrect response. The subject was instructed to decide which of the two responses was last associated with the stimulus, and to guess between the two if necessary.

(c) Four-alternative forced-choice (FC-4): The situation was the same as for FC-2, except that the correct response was displayed along with three randomly-chosen distractors. All four responses appeared on one horizontal line, separated from one another by about $1\frac{1}{2}$ in. Again, the position of the correct response among the four was randomly determined.

(d) Twenty-six alternative forced-choice (FC-26): Rather than a display of all the letters of the alphabet, the word "alphabet" appeared on the CRT below and to the right of the stimulus. The subject was instructed to try to remember the response last associated with the stimulus, and to guess if necessary. (2) Once the appropriate typewriter key was depressed, feedback was given for $1/2$ sec. by the messages "correct" or "wrong" displayed in the middle of the CRT. The word study then appeared on the upper face of the CRT. Beneath the word study a stimulus-response pair appeared. The stimulus was the same one used in the preceding test portion of the trial. The response was randomly selected from the letters of the alphabet, with the single restriction that it be different from the one just tested in the present trial. The study period lasted 2 sec., and was followed by a CRT black-out for a $1/2$ sec. intertrial interval.

RESULTS

In previous experiments using the continuous task employed here (Atkinson, Brelsford, and Shiffrin, 1967; Brelsford and Atkinson, 1968; Brelsford, Shiffrin, and Atkinson, 1968) it was observed that a slight

warm-up takes place at the start of each session. For this reason the data from the first 25 trials of each session, and from the entire first session, were excluded from all analyses. Figure 1 presents the probability of a correct response as a function of lag for the forced-choice conditions in the lower panel; hit and false alarm probabilities are presented in the upper panel. It is clear that as the number of alternatives in the forced-choice task increases, the overall performance decreases. It can be seen that the curves drop sharply at first and slowly thereafter for the forced-choice conditions. The hit curve decreases in a similar manner, while the false alarms show a slow trend upward over lags. It is noteworthy that by lag 14 the forced-choice curves are still falling gradually and have not reached the chance level; chance performance should be $1/2$, $1/4$, $1/26$ for the FC-2, FC-4, FC-26 conditions, respectively. We should note that the procedure used to select a stimulus for test results in a geometrically decreasing number of observations at successive points on the lag curve. For each of the four types of tests, the number of observations ranges from approximately 1000 at lag 1 to 100 at lag 14.

Figures 2 and 3 present the latencies of responses for the four conditions, conditionalized upon a correct response (Fig. 2) and upon an error (Fig. 3). A salient feature of these data is that reaction time is not a monotonic function of the number of response alternatives. The subjects required less time to respond to twenty-six alternatives than to four. This is true for both correct and incorrect responses. It is also clear that the latencies appear to be relatively independent of the lag of the item, maintaining rather constant values across all lags

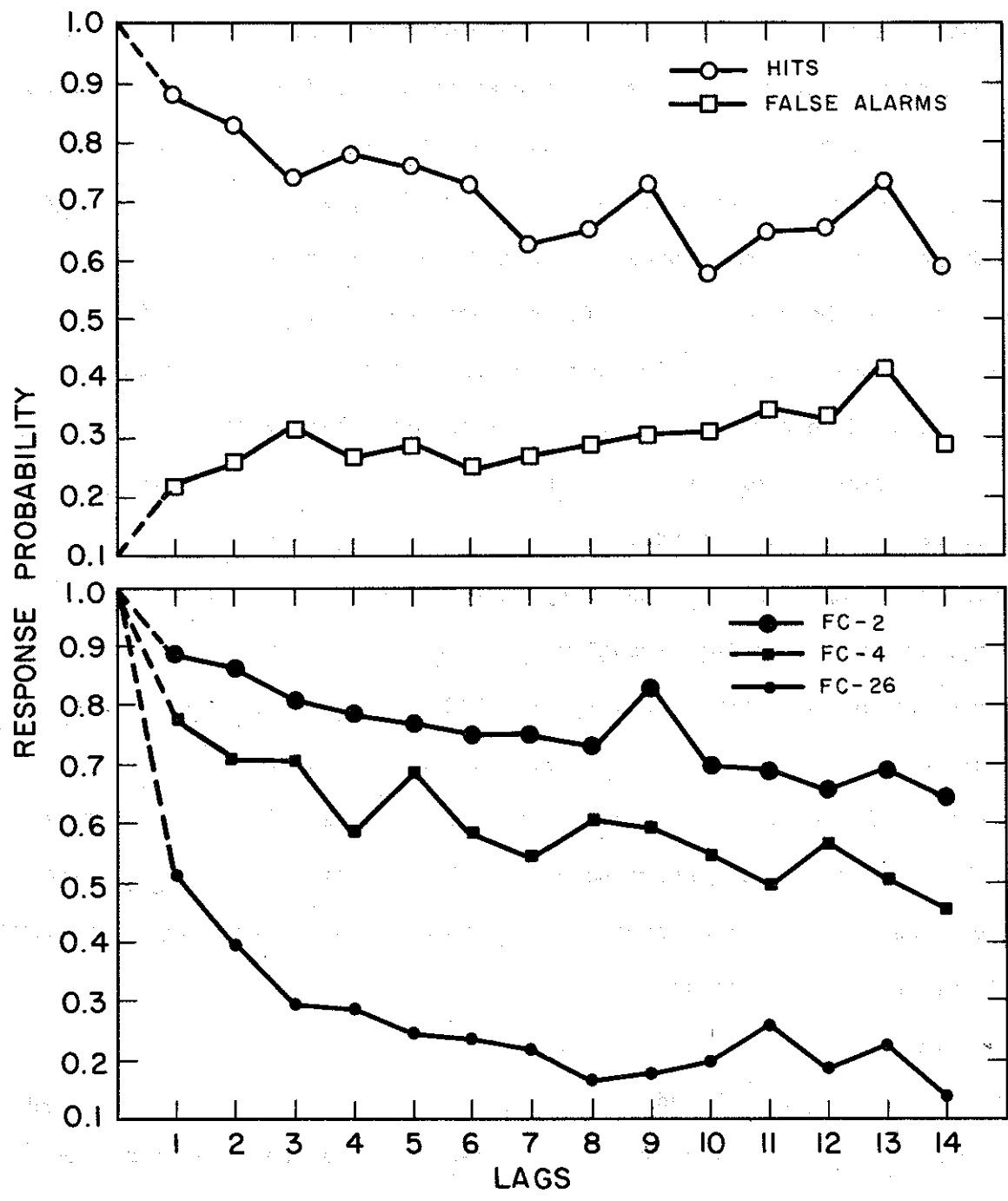


Fig. 1. Observed response probabilities as a function of lag.

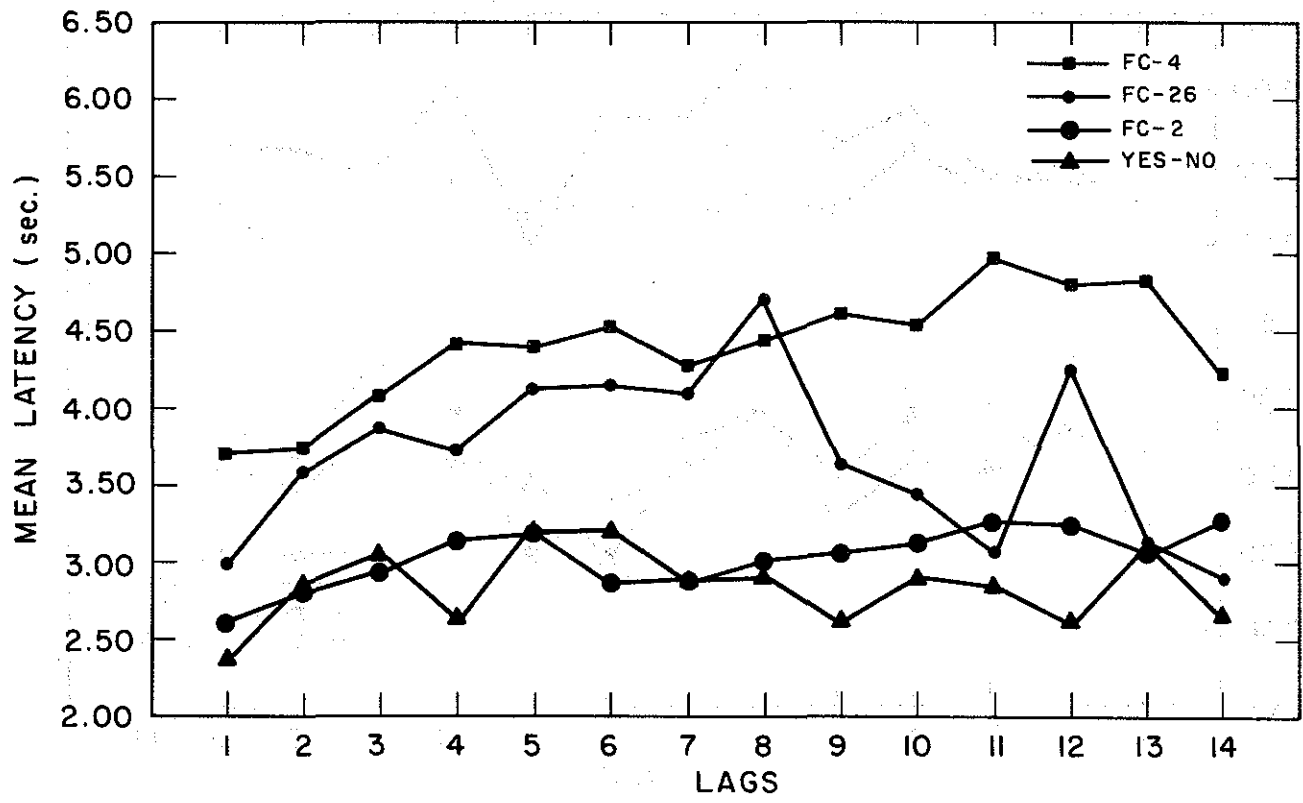


Fig. 2. Latency given a correct response as a function of lag for four types of tests.

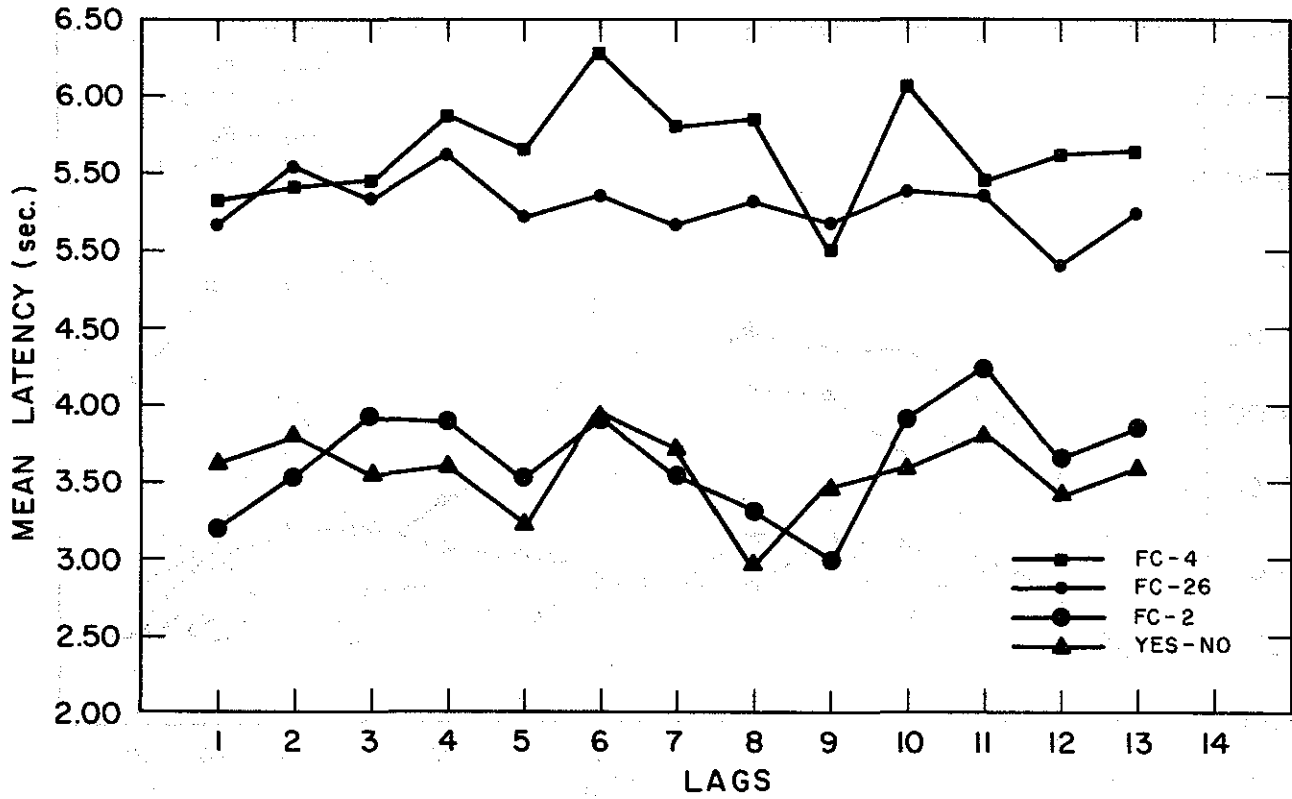


Fig. 3. Latency given an error as a function of lag for four types of tests.

plotted. The coincidence of the FC-2 and yes-no curves is also noteworthy, especially since both of these tests offered a 50% opportunity for a chance success.

MODEL PREDICTIONS

For both Model I and Model II three different cases are considered. These cases represent increasing reliance upon the rehearsal buffer for correct responding.

Case 0. Here the buffer is eliminated entirely. This is accomplished by setting α (the probability that an item enters the buffer) equal to zero. This is to be interpreted to mean that an item is studied and transfers information at that time of study into LTS. All responses are made on the basis of the information retrieved from LTS.

Case 1. Limited reliance on the buffer is achieved by setting r (size of buffer) at 1. Most responding is based on retrieval of information from LTS, but there is some likelihood of correct performance due to an item's presence in the buffer at the time of test.

Case 2. Finally, complete freedom for all parameters allows the full power of the model to be realized. It should be clear that if r estimates to be 1, then this case reduces to Case 1. For this reason we shall specify that in tests of Case 2, r is 2 or greater.

In order to estimate the parameters of these models and evaluate goodness-of-fit, the following function is defined:

$$\chi^2(\alpha, r, \theta, \tau, c) = \sum_i \left\{ \frac{1}{N_i P_i} + \frac{1}{N_i - N_i P_i} \right\} \left\{ N_i P_i - O_i \right\}^2 \quad (9)$$

where the sum is over all 14 lags. Here P_i denotes the theoretical

probability of a response and N_i is the total number of observations at the i^{th} point. For the forced-choice conditions O_i is the observed number of correct responses for the i^{th} point, whereas for the yes-no condition O_i is the observed number of hits or false alarms. The more accurate the model, the closer O_i should be to $N_i P_i$ and the smaller the χ^2 . An overall χ^2 value was obtained by minimizing the function simultaneously over the five curves (three forced-choice curves, the hits, and false alarms) displayed in Figure 1. This minimum χ^2 is based upon $14 \times 5 = 70$ degrees of freedom, minus one for each parameter estimated. Although not a true χ^2 , this function has been shown to be adequate for the parameter estimation technique employed here. The minimization was performed on a computer as follows. Tentative values of the parameters are selected, the associated χ^2 values computed, and this is repeated for another set of parameter values until the space of possible values has been systematically explored. The computer program yields the parameter values which provide the minimum χ^2 . When enough points in the parameter space are scanned, the method yields a close approximation to the analytic solution. A further discussion of this procedure may be found in Atkinson, Bower, and Crothers (1965).

Table 1 presents the parameter values corresponding to the best fits for each of the three cases of Model I and Model II. The χ^2 corresponding to each of these fits is given in Table 2, divided into the components contributed by each of the five test conditions to the total χ^2 . It is clear that Model II provides considerably better fits to the lag data than does Model I. The χ^2 of the best-fitting case of the latter model are approximately four times those of Model II. Case 1

Table 1

Best fit parameters for Models I and II (Cases 0, 1 and 2)

		α	r	θ	τ	γ
Model I	Case I	0*	-	0.80	0.86	0.53
	Case II	0.50	1*	0.42	0.90	0.53
	Case III	0.59	2	0.17	0.92	0.53
		α	r	θ	τ	c
Model II	Case I	0*	-	1.86	0.93	0.54
	Case II	0.75	1*	0.86	0.93	0.54
	Case III	1.00	2	0.45	0.95	0.54

*Parameter not estimated in grid search.

Table 2

Chi-square by conditions for Models I and II (Cases 0, 1 and 2)

		Hits	False Alarms	FC-2	FC-4	FC-26	Total
Model I	Case I	42	28	227	206	428	931
	Case II	43	23	220	193	435	914
	Case III	40	30	233	214	363	880
Model II	Case I	35	27	46	48	82	238
	Case II	25	12	35	36	55	163
	Case III	28	53	25	30	83	219

of Model II, with a limited reliance on the buffer, provides the best fit of all. These fits are depicted in the following figures: Figure 4 presents the lag curves for the best-fitting case of Model I (Case 2). Cases 0, 1, and 2 of Model II are displayed in Figures 5, 6, and 7, respectively. In each figure the lower panel depicts the three forced-choice conditions and the upper panel shows the hit and false alarm curves. Inspection of Figure 4 reveals poor fits for Model I in all forced-choice conditions: the curves are predicted to be too close together. Model II, on the other hand, provides a good fit to all three forced-choice curves. The shape of the false alarm curve is quite good for Case 1 in particular, as the data of Table 1 suggest.

A number of aspects of the parameter values for the three cases should be noted. It was found that the value of c (the criterion value) was invariant over the three cases of Model II. The model predicts such a constant criterion value on the following basis. The parameter c affects only the hit and false alarm curves, and is primarily determined by the asymptote of the false alarm curve. When longer lags are examined, and thus the contribution of the buffer is minimal, the false alarm predictions are essentially determined by the normal deviate corresponding to the value of c that predicts the observed proportion of false alarms. From the normal distribution, then, it is found that the value of $c = .54$ corresponds to a false-alarm asymptote of approximately 0.29.

The value of θ varies widely across the cases considered, decreasing with increasing reliance on the buffer as we progress from Case 0 to Case 2. This sensitivity of θ to the importance of the buffer follows directly from the general model, and it may be instructive to present

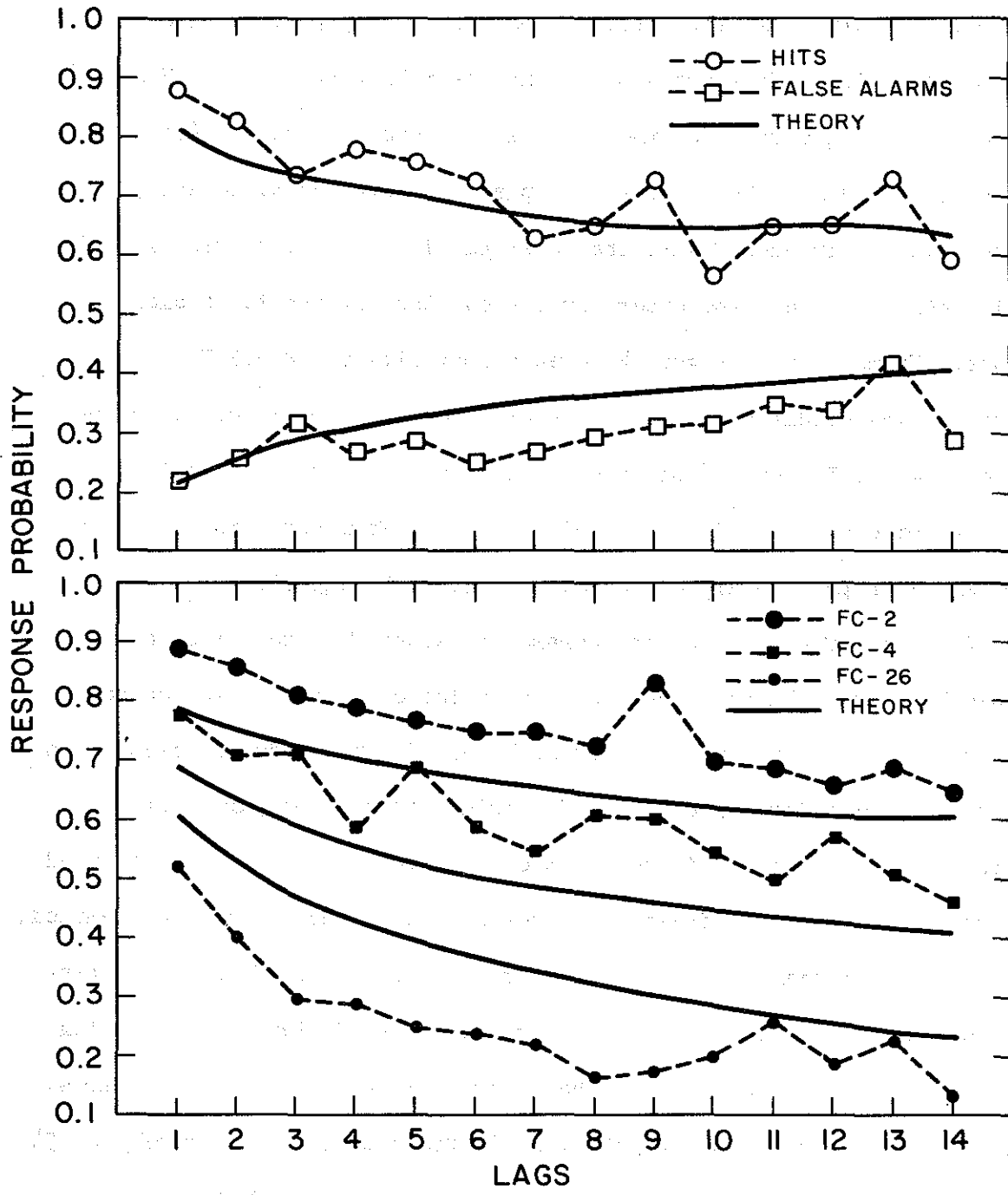


Fig. 4. Observed and predicted response probabilities as a function of lag for Model I, Case 2.

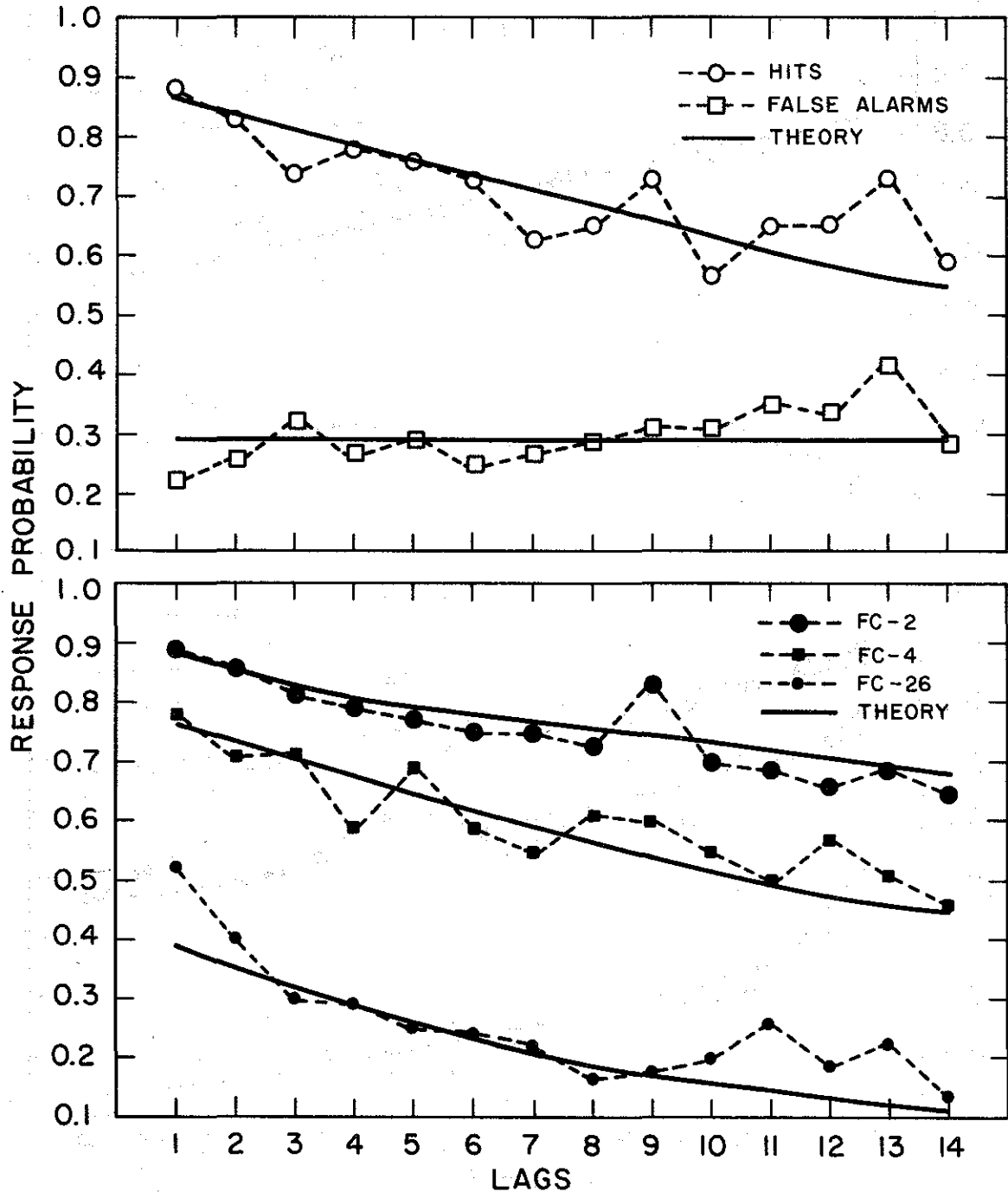


Fig. 5. Observed and predicted response probabilities as a function of lag for Model II, Case 0.

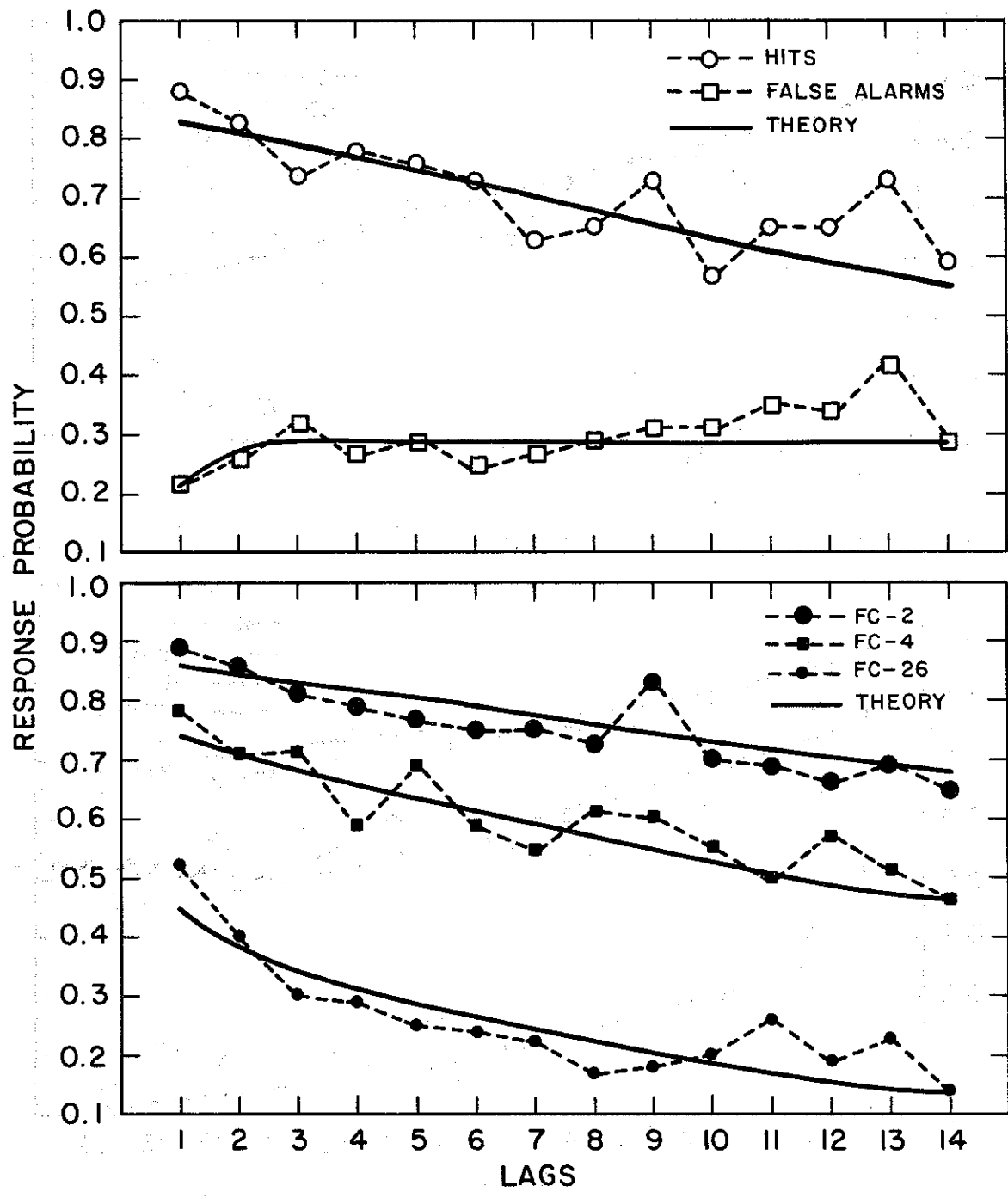


Fig. 6. Observed and predicted response probabilities as a function of lag for Model II, Case 1.

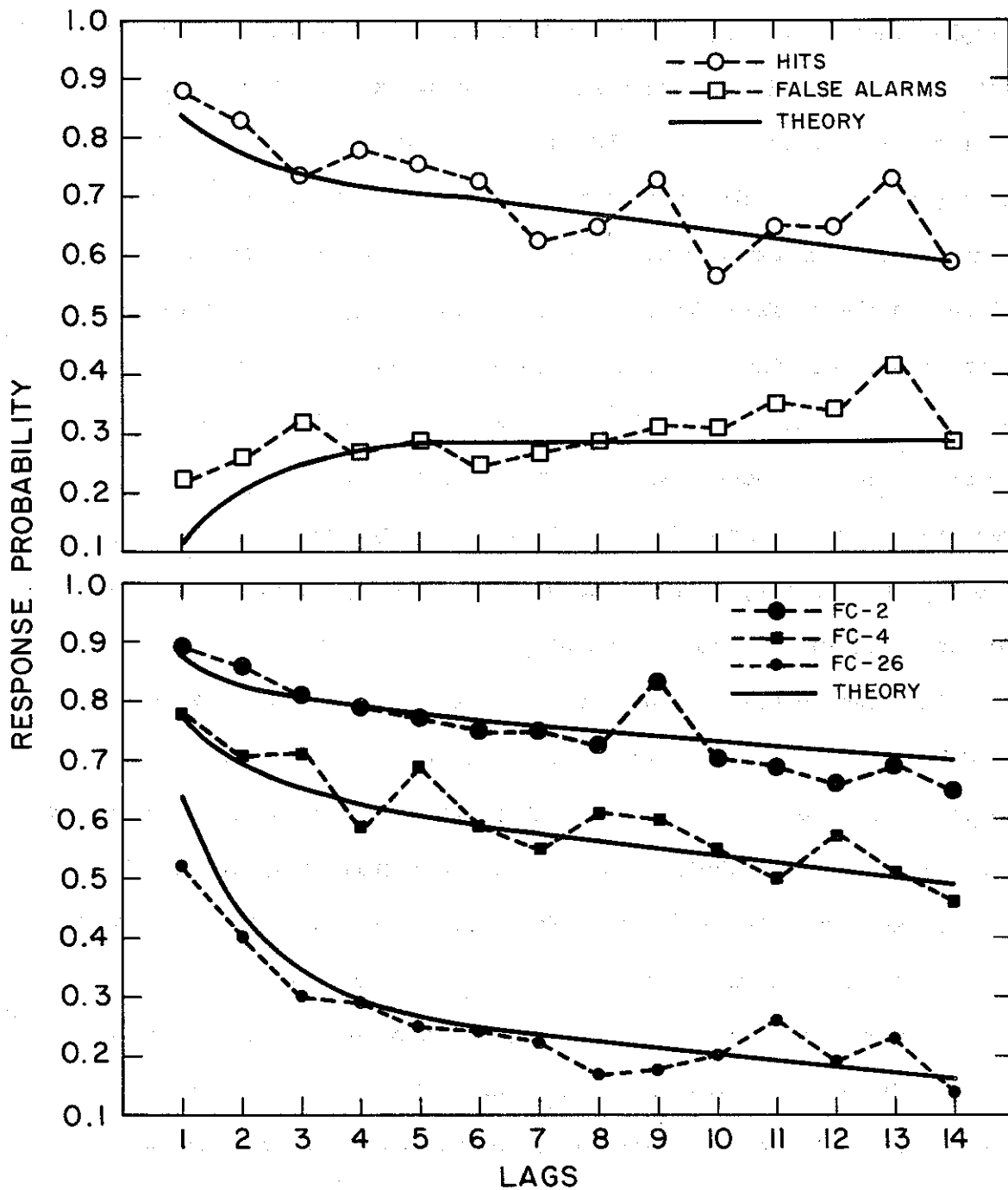


Fig. 7. Observed and predicted response probabilities as a function of lag for Model II, Case 2.

the explanation explicitly. When there is only a one-time transference of information to LTS (α being zero), θ must be very high in order that sufficient information exist in LTS for the subject to produce any correct responses. Note that θ is not a probability, but rather represents the amount of information being transferred to LTS during the item's residence in the buffer. When there is a relatively high probability that an item enters the buffer as in Case 1, θ need not be so high; and with a buffer of size 2 as in Case 2, θ is lower still. Two processes are responsible for these changes in the transfer rate: relatively more correct responses are made from the buffer (as opposed to retrievals from LTS), and further, an item has more opportunity to remain in the buffer and build up information in LTS, the larger the buffer.

In the cases where α is permitted to vary, the estimates were quite high: 0.75 with a buffer of size 1, and 1.00 with a buffer of size 2. These values indicate a relatively high probability of new items entering the buffer, and are at variance with previous results (e.g., Atkinson, Brelsford, and Shiffrin, 1967; Brelsford, Shiffrin, and Atkinson, 1968) where the values of α were usually much lower. A possible reason for the high α values arises from the fact that there were no tests given at lag 0. Previous studies have established that the probability of a correct response at lag 0 is virtually unity, and therefore such tests were eliminated here in order to collect more data at other lags. It seems likely, however, that the absence of lag 0 tests may have influenced the rehearsal strategies subjects adopted, and hence the parameter values obtained. Specifically, the general model postulates that the probability of a correct response at lag 0 is unity, since every item enters

STS. Thus, when lag 0 tests are included, subjects are assured of perfect performance on many items (1/8 of them, for example, when there are eight stimuli in the to-be-remembered set) regardless of whether or not the item is entered into the rehearsal buffer. In the present experiment, however, all tests were at lag 1 or greater, and thus there was little opportunity for correct responding without entering the item into the buffer. An optimal strategy here might well be to devote some effort to the coding of each item, which corresponds to entering each item into the buffer.

Examination of Figures 5, 6, and 7 reveals that the major discrepancies between Model II and the data occur at short lags (i.e., lags 1 to 3) for the FC-26 and false-alarm curves. For Case 0, the probability of a false alarm at any lag should depend only upon the value of c (the criterion value) since all responses reflect retrieval from LTS. False alarms should therefore remain constant over lags, assuming that c does not depend on lag. The observed points, however, show a definite rise over the shorter lags. Moreover, it may be assumed that no false alarms would have occurred at lag zero, since all previous studies obtained perfect performance at lag 0. Thus, the constant probability of a false alarm across lags predicted by Case 0 is clearly incorrect.

It should be clear that the lag curves reflect STS processes primarily at the shorter lags, and LTS processes as the lag increases. Without explicit account of the short-term processes, the TSD analysis applied here in Case 0 of Model II is not adequate to handle the data. Its most obvious weakness is the prediction of a constant false-alarm rate across all lags, whereas data from previous studies has shown virtually perfect

performance at lag 0, and the present study has shown the false-alarm rate rising over early lags toward an asymptote. On the other hand, the general model of previous studies (Model I) fails to provide an adequate fit to any of the five curves, despite its assumption of separate short-term and long-term memory states. The best account of the data utilizes a compromise between the two approaches, with TSD analysis applied to LTS processes while the buffer determines short-term effects.

With regard to the FC-26 condition, Table 2 reveals that the largest component of the total X^2 was contributed by this condition in all cases. The relatively poor performance of Model II for FC-26 may be due to the following considerations. It is unlikely that subjects in the FC-26 test would undertake selection from among all 26 possible alternatives, especially since the alphabet was not actually displayed on the CRT. It is plausible that the subject in a FC-26 test considers only a subset of alternatives. Suggestive of such a limited search is the latency data: the subjects take a shorter time to respond in the FC-26 condition than in the FC-4 condition. If the subjects were in fact considering a random subset of letters, then the lag curves for the FC-26 condition would be displaced downward. As can be seen in Figure 7, this would result in a better fit for the case where r equals 2.

This possibility was explored by postulating a parameter m , which denotes the size of the subset searched by a subject when faced with 26 alternatives. Since a subset of size m has probability $m/26$ of containing the correct response, the probability of a correct response at lag i would be reduced by proportion $m/26$. A search of the parameter space was carried out, and the best fitting values are given in Table 3. The

χ^2 by conditions are shown in Table 4. The only case affected to any degree is Case 2, with some improvement in the fit for the FC-26 data with an estimate of $m = 19$. This does not appear to be a very rewarding channel of investigation, although the latency data suggest that there may be other possibilities along this line which merit consideration.

Table 3

Best fit parameters for extended Model II (Cases 0, 1 and 2)

	α	r	θ	τ	c	M
Case I	0*	-	1.86	0.93	0.54	26
Case II	0.75	1*	0.86	0.93	0.54	25
Case III	0.98	2	0.46	0.95	0.54	19

*Parameter not estimated in grid search.

Table 4

Chi-square by conditions for extended Model II (Cases 0, 1 and 2)

	Hits	False Alarms	FC-2	FC-4	FC-26	Total
Case I	35	27	46	48	82	238
Case II	25	12	35	36	54	162
Case III	25	53	29	27	70	204

APPENDIX

Let $1 - \beta_0^{(i)}$ be the probability that the i^{th} item enters the buffer. To find $1 - \beta_0^{(i)}$, we conditionalize on whether or not the previous item entered the buffer:

$$1 - \beta_0^{(i)} = \Pr(i^{\text{th}} \text{ enters} \mid (i-1)^{\text{st}} \text{ did not enter}) + \Pr(i^{\text{th}} \text{ enters} \mid (i-1)^{\text{st}} \text{ did enter}) . \quad (\text{A1})$$

Now taking each of these terms separately, the probability that the i^{th} item enters given that the $(i-1)^{\text{st}}$ entered, is the probability that the i^{th} item was already in the buffer plus the probability that the i^{th} item was not there and enters given that the $(i-1)^{\text{st}}$ entered, or

$$\left[\frac{r-1}{s-1} + \left(1 - \frac{r-1}{s-1}\right)\alpha \right] [1 - \beta_0^{(i-1)}] , \quad (\text{A1a})$$

where s is the size of the stimulus set ($s = 8$ in present experiment).

Similarly, the probability that the i^{th} item enters given that the $(i-1)^{\text{st}}$ did not equals the probability that the i^{th} item was already in the buffer plus the probability that it was not there and enters, given that the $(i-1)^{\text{st}}$ did not enter, or

$$\left[\frac{r}{s-1} + \left(1 - \frac{r}{s-1}\right)\alpha \right] \beta_0^{(i-1)} . \quad (\text{A1b})$$

Thus Eq. A1 becomes:

$$1 - \beta_0^{(i)} = \left[\frac{r-1}{s-1} + \left(1 - \frac{r-1}{s-1}\right)\alpha \right] [1 - \beta_0^{(i-1)}] + \left[\frac{r}{s-1} + \left(1 - \frac{r}{s-1}\right)\alpha \right] \beta_0^{(i-1)} . \quad (\text{A2})$$

Now we further assume that the probability that any item enters the buffer is the same for every item. Therefore, we may drop the superscripts in Eq. A2 and write

$$\begin{aligned}
1 - \beta_0 &= \left[\frac{r-1}{s-1} + \left(1 - \frac{r-1}{s-1}\right)\alpha \right] [1 - \beta_0] \\
&+ \left[\frac{r}{s-1} + \left(1 - \frac{r}{s-1}\right)\alpha \right] \beta_0 .
\end{aligned}
\tag{A3}$$

Collecting terms and simplifying,

$$\beta_0 = \frac{(s-r)(1-\alpha)}{s-\alpha} .$$

Having obtained β_0 , we can easily find β_1 . It is simply the joint probability that the item enters, a new item (N-item) occurs on the next trial, and this new item enters and knocks out the item in question:

$$\beta_1 = (1 - \beta_0) \left(1 - \frac{r-1}{s-1}\right) \alpha \frac{1}{r} .$$

We shall now consider the rest of the β_i terms. Let ϕ represent the item of interest, i.e., the item tested at lag i . Let ρ represent the previous item, i.e., the item presented on the previous trial. Again, as with β_0 , we will conditionalize on whether or not ρ got into the buffer or not. We also conditionalize on ϕ still being in the buffer (i.e., not yet lost) at lag i . Define $q_a^{(i)}$ and $q_b^{(i)}$ as follows:

$$q_a^{(i)} = \Pr(\rho \text{ is in buffer and } \phi \text{ not yet lost by lag } i)$$

$$q_b^{(i)} = \Pr(\rho \text{ is not in buffer and } \phi \text{ not yet lost by lag } i)$$

Then

$$\beta_2 = (1-\beta_0) \left[q_a^{(1)} \left(\frac{s-r}{s-2} \right) \frac{\alpha}{r} + q_b^{(1)} \left(\frac{s-r-1}{s-2} \right) \frac{\alpha}{r} \right]$$

$$\beta_3 = (1-\beta_0) \left[q_a^{(2)} \left(\frac{s-r}{s-2} \right) \frac{\alpha}{r} + q_b^{(2)} \left(\frac{s-r-1}{s-2} \right) \frac{\alpha}{r} \right]$$

and in general

$$\beta_i = (1-\beta_0) \left[q_a^{(i-1)} \left(\frac{s-r}{s-2} \right) \frac{\alpha}{r} + q_b^{(i-1)} \left(\frac{s-r-1}{s-2} \right) \frac{\alpha}{r} \right] \quad (A4)$$

The two terms in the brackets represent the conditionalizing upon whether the previous item (ρ) did or did not get into the buffer, given that the item of interest (ϕ) is still there. If the previous item did get in the buffer (with probability $q_a^{(i-1)}$), then with probability $\frac{s-r}{s-2}$ a new item is tested, enters the buffer (with probability α) and knocks out ϕ (with probability $\frac{1}{r}$). If the previous item did not get in the buffer, then with probability $(s-r-1)/(s-2)$ a new item is tested, enters the buffer (with probability α) and knocks out ϕ (with probability $\frac{1}{r}$). The denominator, $s-2$, comes from the fact that we consider all remaining items other than ρ and ϕ . The numerators, $s-r$ and $s-r-1$, are the sets from which we choose a new item, given all our conditionalizations.

The problem now remains to derive the $q_a^{(i)}$ and $q_b^{(i)}$ terms. For this we use the following transition matrix Q on the $q_a^{(i)}$ and $q_b^{(i)}$ terms:

$$Q = \begin{array}{cc} & \begin{array}{c} q_a^{(i+1)} \\ q_b^{(i+1)} \end{array} \\ \begin{array}{c} q_a^{(i)} \\ q_b^{(i)} \end{array} & \left[\begin{array}{cc} \frac{r-2}{s-2} + \left(1 - \frac{r-2}{s-2}\right) \alpha \left(\frac{r-1}{r}\right) & \left(1 - \frac{r-2}{s-2}\right) (1-\alpha) \\ \frac{r-1}{s-2} + \left(1 - \frac{r-1}{s-2}\right) \alpha \left(\frac{r-1}{r}\right) & \left(1 - \frac{r-1}{s-2}\right) (1-\alpha) \end{array} \right] \end{array} \quad (A5)$$

For example, the $q_a^{(i)}$ to $q_a^{(i+1)}$ transition represents the probability that an item ρ in the buffer on one trial will be there on the next trial. This is simple the probability that some other item in the buffer is tested (and therefore replaces itself leaving ρ and ϕ undisturbed) plus the probability that an item not currently in the buffer is tested,

enters the buffer (with probability α) but does not knock out ρ . The other transitions are derived similarly. The start vector is simply:

$$S_1 = [q_a^{(1)}, q_b^{(1)}] = \left[\frac{r-1}{s-1} + \left(1 - \frac{r-1}{s-1}\right)\alpha\left(\frac{r-1}{r}\right), \left(1 - \frac{r-1}{s-1}\right)(1-\alpha) \right] \quad (A6)$$

and so to generate the q 's we have:

$$[q_a^{(i)}, q_b^{(i)}] = S_1 Q^{i-1}.$$

References

- Atkinson, R. C., Bower, G. H., and Crothers, E. J. An Introduction to Mathematical Learning Theory. New York: John Wiley & Sons, 1965, 429 pp.
- Atkinson, R. C., and Shiffrin, R. M. Mathematical models for memory and learning. Technical Report 79, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965. (To be published in Readiness to Learn: Proceedings of the Third International Interdisciplinary Conference on Learning, Remembering and Forgetting, edited by D. P. Kimble, New York: The New York Academy of Sciences.)
- Atkinson, R. C., and Shiffrin, R. M. Human memory: A proposed system and its control processes. In The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2, edited by K. W. Spence and J. T. Spence. New York: Academic Press, 1968a, 89-195.
- Atkinson, R. C., and Shiffrin, R. M. Some speculations on storage and retrieval processes in long-term memory. Technical Report 127, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1968b.
- Atkinson, R. C., Brelsford, J. W., Jr., and Shiffrin, R. M. Multi-process models for memory with applications to a continuous presentation task. Journal of Mathematical Psychology, 1967, 4, 277-300.
- Brelsford, J. W., Jr., and Atkinson, R. C. Recall of paired-associates as a function of overt and covert rehearsal procedures. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 730-736.

- Brelsford, J. W., Jr., Shiffrin, R. M., and Atkinson, R. C. Multiple reinforcement effects in short-term memory. British Journal of Mathematical and Statistical Psychology, 1968, in press.
- Bernbach, H. A. Decision processes in memory. Psychological Review, 1967, 74, 462-480.
- Donaldson, W., and Murdock, B. B., Jr. Criterion change in continuous recognition memory. Journal of Experimental Psychology, 76, 1968, 325-330.
- Kintsch, W. Memory and decision aspects of recognition learning. Psychological Review, 1967, 74, 496-504.
- Swets, J. A. Signal Detection and Recognition by Human Observers: Contemporary Readings. (Ed.) New York: John Wiley & Sons, 1964.
- Wickelgren, W. A., and Norman, D. A. Strength models and serial position in short-term recognition memory. Journal of Mathematical Psychology, 1966, 3, 316-347.