A THREE STATE MARKOV MODEL FOR LEARNING

by

John Theios

TECHNICAL REPORT NO. 40

September 22, 1961

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Applied Mathematics and Statistics Laboratories
STANFORD UNIVERSITY
Stanford, California

A THREE STATE MARKOV MODEL FOR LEARNING[1]

John Theios

Stanford University

In many learning situations, the response under study initially has zero probability of occurring, but asymptotically, the response probability approaches unity. A few situations of this type are instrumental avoidance conditioning, classical defense and appetitive conditioning, and reversal learning. These situations will be referred to as simple conditioning (SC). In the past, SC has been analyzed in terms of response strength or linear operator models (Hull, 1943; Estes, 1950; Bush & Mosteller, 1955), which assume that the strength or probability of a learned response increases gradually during the course of learning. Recently, it has been found that Markov models, which assume that learning takes place on single trials in an all-or-none fashion, more adequately describe some types of verbal learning than do the linear models (Bower, 1960; Estes, 1960). It is quite possible that SC is also characterized by some sort of discrete learning as opposed to gradual learning. This possibility is further enhanced by the fact that the "zero to unity" response probabilities, characteristic of SC, should lend themselves nicely to discrete conditioning states, which a Markov interpretation would require (cf., Suppes & Atkinson, 1960). The present paper presents

1/ I am indebted to Dr. Gordon H. Bower for deriving a number of the theoretical predictions and for his valuable interest and advice during the development of this paper.

a three state absorbing Markov model for SC, and then compares the theoretical predictions to actual data collected in an extensive experiment on avoidance conditioning of rats.

THE TWO-ELEMENT STIMULUS SAMPLING MODEL

The basic learning theory assumed in this paper is that proposed by Suppes and Atkinson (1960). They assume that any learning situation can be represented by a finite set of stimulus patterns and that each component pattern of the stimulus set is conditioned to exactly one response from the set of mutually exclusive responses available to the subject. On each trial only one pattern is sampled, and the subject makes that response to which the sampled pattern is conditioned. If the response that occurs is not correct in relation to the reinforcement that occurs, it is assumed that the reinforcement elicits or forces the correct response, which becomes conditioned to the sampled pattern with a fixed probability, c.

In SC only one response is reinforced on every trial. The particular model to be proposed assumes that SC situations can be represented by exactly two stimulus patterns and that on any given trial each pattern has probability 1/2 of being sampled by the subject. The fundamental axioms of the model are as follows:

Identification Axioms

I1. A simple conditioning situation may be represented by exactly two stimulus patterns.

I2. At the start of conditioning, neither of the two patterns is conditioned to the correct, A_1, response.

Conditioning Axioms

C1. On every trial, each of the two stimulus patterns is conditioned to exactly one response.

C2. The stimulus pattern that is sampled on a trial becomes conditioned to the reinforced response with a fixed probability, c. If the pattern is already conditioned to the reinforced response, it remains so conditioned.

C3. The stimulus pattern not sampled on a given trial does not change in conditioning on that trial.

C4. The probability, c, that the sampled pattern will become conditioned to the reinforced response is independent of the trial number and preceding sequence of events.

Sampling Axioms

S1. Exactly one pattern is sampled on each trial.

S2. Each of the two patterns has probability $1/2$ of being sampled on a given trial.

S3. On any trial, the probability of sampling a given pattern is independent of the trial number and preceding sequence of events.

Response Axiom

R1. On any trial, that response is made to which the sampled pattern is conditioned.

According to the axioms, the learning process in SC may be described as an absorbing Markov process with three states. At the start of the experiment, when neither of the two stimulus patterns is conditioned to the correct response, the process is in conditioning state S_0. After one of the patterns becomes conditioned to the correct response, the process is in conditioning state S_1, where the probability of a correct response is 1/2. Finally, when both patterns are conditioned to the correct response, the process is in the absorbing state, S_2, where the probability of a correct response is unity. The trees of the Markov process are given in Figure 1. The matrix of transition probabilities, P, in canonical form is

$$(1) \qquad P = \begin{array}{c|c|cc} & S_2 & S_1 & S_0 \\ \hline S_2 & 1 & 0 & 0 \\ \hline S_1 & \frac{c}{2} & 1-\frac{c}{2} & 0 \\ S_0 & 0 & c & 1-c \end{array} = \begin{array}{|c|c} I & 0 \\ \hline R & Q \end{array} .$$

The fundamental matrix of the chain, N, is

$$(2) \qquad N = \sum_{k=1}^{\infty} Q^k = (I - Q)^{-1} = \begin{array}{c|cc} & S_1 & S_0 \\ \hline S_1 & \frac{2}{c} & 0 \\ S_0 & \frac{2}{c} & \frac{1}{c} \end{array} ,$$

where I is the identity matrix and Q is the matrix of the transient states. The entries n_{ii} of the matrix N give the mean of the total number of times the process will be in transient state S_i (for $i=0,1$).

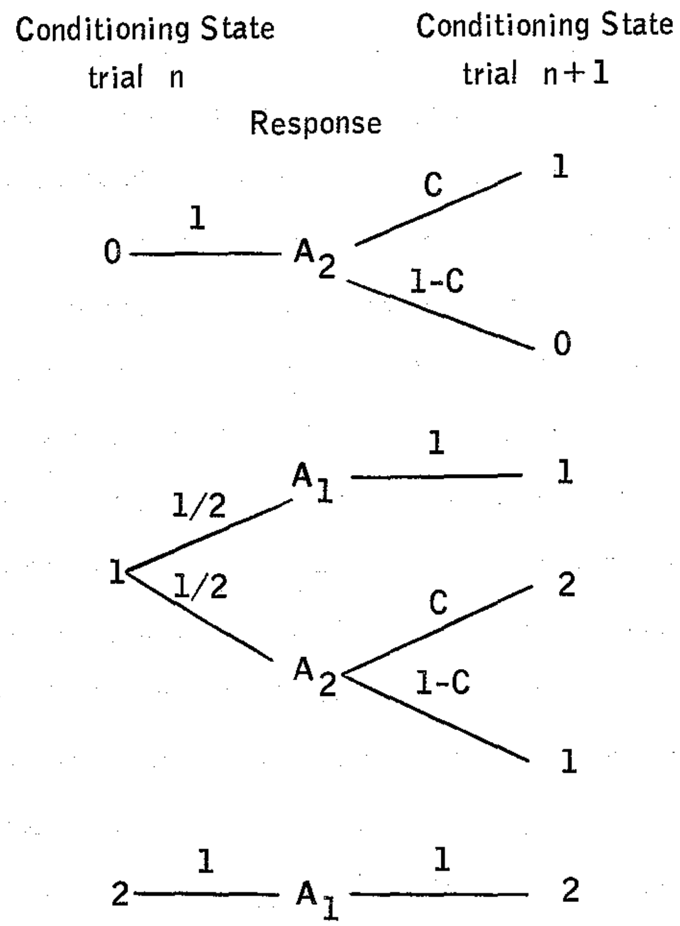

Fig. 1. The trees of the Markov process.

PREDICTIONS AND DATA

A large number of predictions which follow from the model will be derived in this section. The derivations for a number of these predictions were obtained by Gordon Bower. As a test of the adequacy of the model, the predictions will be compared to data collected in an avoidance learning experiment where 50 rats served as subjects. The apparatus was a modified Miller-Mowrer electric shock box. The correct response (A_1) was to run from one compartment to the other within 3 seconds after a buzzer sounded and a light came on in the white compartment. Special care was taken to reduce the stimulus situation drastically so that the situation could be represented by only two stimulus patterns. The reduction was achieved, for example, by reducing the external distractions for the subject, using a high intensity shock (255 volts), running the subjects only one way (e.g., always black to white) rather than having them shuttle, and giving all trials in one experimental session at 20 second intertrial intervals. The procedure was to place the subject in one compartment and turn on the buzzer and light as the door between the compartments was opened. If the subject did not run into the other compartment within 3 seconds, the rat was shocked until he escaped into the other compartment. The buzzer, light, and shock terminated when the other compartment was entered. After 20 seconds the subject was returned to the first compartment, and another trial was given. A subject was run until he met a criterion of 20 consecutive successful avoidance responses. When a subject met the criterion he was given reversal learning (e.g., if he had originally learned to run from black to white, during reversal learning, he learned to run from white to black). Since there were no important differences between the data of original and reversal learning, the data from the two series were pooled, yielding 100 response sequences to test the model.

Bernoulli Properties of the Model

One cannot observe on what trial a transition from conditioning states S_0 to S_1 occurs. However, if there are some trials between the first success and the last error, we can be sure that the subject is in state S_1 on these trials. For surely, if the subject has made one success, at least one of the two stimulus patterns is conditioned to the A_1 response; and if on a later trial the subject makes an error, then at least one of the patterns is not conditioned to the A_1 response. Since deconditioning does not occur in the present model, the above two patterns must be distinct, and by definition, the subject must be in conditioning state S_1. According to the model, the probability of a success in state S_1 is a constant, 1/2. Also, the conditional probability of a success on trial n+1 given a success on trial n is a constant, 1/2, since conditioning cannot occur on success trials:

$$(3) \qquad P(A_{1,n+1} | A_{1,n} \cap S_{1,n}) = \frac{1}{2} .$$

Thus, according to the model, the sequence of responses between the first success and the last error should be an independent Bernoulli sequence with $p = q = \frac{1}{2}$, and all statistics relevant to coin flipping experiments should be applicable to the response sequences during these trials.

For example, defining h_j as the probability of obtaining a run of successes j trials long between the first success and the last error, the probability distribution of h_j will be equal to the expected probability distribution of obtaining a run of heads in a coin flipping experiment, which is

$$(4) \qquad P(h_j) = p^{j-1}q = (\tfrac{1}{2})^j, \qquad \text{For } j=1,2,3, \ldots .$$

The mean and variance of the distribution of h_j will be

$$(5) \qquad E(\bar{h}) = \sum_{j=1}^{\infty} j(\tfrac{1}{2})^j = 2 ,$$

and

$$(6) \qquad \text{Var}(\bar{h}) = \sum_{j=1}^{\infty} (j-2)^2 (\tfrac{1}{2})^j = 2 .$$

The obtained and predicted expectation and standard deviation of the mean length of runs of successes between the first success and the last error are given in Table 3.

According to the linear models, as n increases, the conditional probability of a success on trial n following a success on trial $n-1$ should also increase. But, the two-pattern sampling model predicts that on trials between the first success and the last error, the conditional probability of a success on trial n given a success on trial $n-1$ should have a constant value of $1/2$ on these trials. The obtained conditional probabilities, given in Figure 2, are approximating a constant value near 1/2, rather than increasing with trials as the linear model would predict. This relationship is, by far, the strongest evidence for the two-pattern sampling model.

Another exacting test of the Bernoulli property of the model is that the response sequences during the trials between the first success and the last error should satisfy the binomial distribution. To provide this test, the data were divided into blocks of four trials, and the number of successes in each block was counted. This sum can take on the values 0, 1, 2, 3, or 4. If the model fits the data, the obtained frequency distribution should not differ significantly from what would be expected from performing a large number of coin flipping experiments in which a coin was tossed four times in each experiment and the distribution of the number of heads in each experiment was tabulated. The obtained and predicted frequencies of successes are given in Table 1. A test of goodness of fit yielded a Chi Square of 1.47, which, with 4 degrees of freedom, indicates that the predictions fit the obtained data very well.

If sampling of the two elements is random, then the outcomes of trials between the first success and the last error should be statistically independent (a zero-order Markov process). This hypothesis can be tested against the alternative hypothesis that the process is a first-order Markov chain. (cf., Suppes and Atkinson, 1960). The matrix of response transitions during the trials between the first success and the last error are given in Table 2. The obtained Chi Square of .07 with one degree of freedom indicates that we cannot reject the hypothesis that the response sequence between the first success and the last error is a zero-order Markov chain.

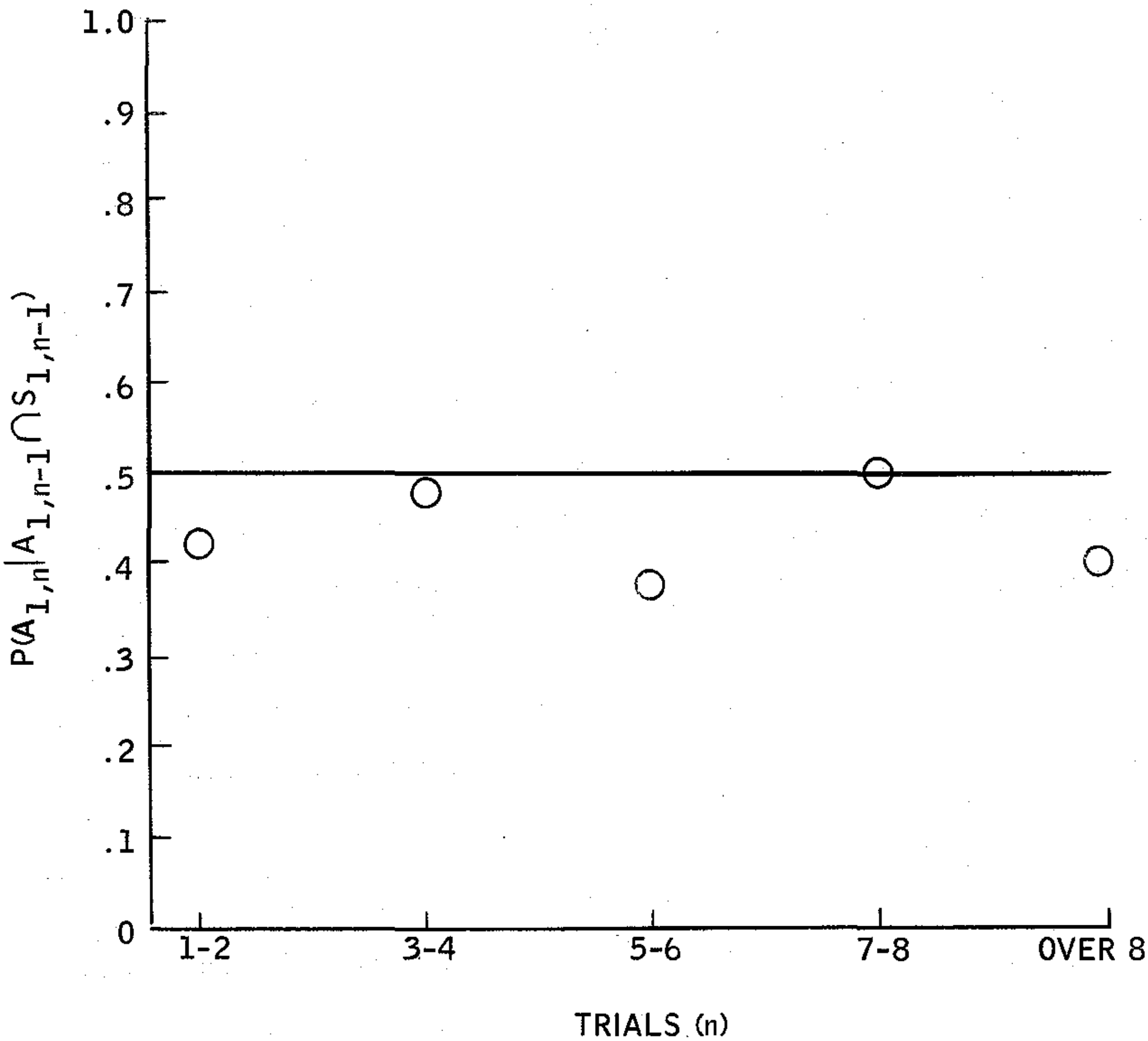

Fig. 2. Conditional Probability of a Success on Trial n Given a Success on Trial n - 1 on Trials Between the First Success and the Last Error.

Table 1

Number of Successes in Blocks of Four Binomial Trials Between the First Success and the Last Error

Number of successes	Obtained frequency	Predicted frequency
0	2	3.1
1	12	12.5
2	17	18.5
3	15	12.5
4	4	3.1
Total	50	50.0

Chi Square = 1.47, df = 4 .

Table 2

Response Transitions Between the First Success and Last Error to Test the Hypothesis of a Zero-Order Markov Process

Trial n	Trial n+1 Success	Trial n+1 Error
Success	90	118
Error	44	54

Chi Square = .07, df = 1 .

Total Errors

Suppose we let t_i represent the number of errors made in transient state S_i. By the axioms, t_i has the geometric distribution given by

(7) $$P(t_i=j) = c(1-c)^{j-1} ,$$

with mean and variance

(8) $$E(t_i) = \frac{1}{c} , \qquad Var\ (t_i) = \frac{(1-c)}{c^2} .$$

In the general case of N stimulus patterns, there will be N transient states. We let T represent the total number of errors before absorption, i.e.,

(9) $$T = t_0 + t_1 + t_2 + \ldots + t_{N-1} .$$

The variable T is the sum of N independent, identically distributed random variables. The probability that T takes on an arbitrary value k is given by the negative binomial distribution

(10) $$P(T=k) = \begin{cases} 0 & \text{for } k < N \\ \binom{k-1}{N-1} c^N(1-c)^{k-N} & \text{for } k \geq N \end{cases} ,$$

which has mean and variance

(11) $$E(T) = \frac{N}{c} , \qquad Var\ (T) = \frac{N(1-c)}{c^2} .$$

It should be noted from Equation 10 that the stimulus sampling models of the type we have been considering make the very strong prediction that the number of errors in any learning sequence must be equal to or larger than the number of stimulus patterns representing the situation. This prediction follows from the assumption that conditioning can occur only when an error has been made.

The expected total errors can serve as a stable estimator for the model's single parameter, c. In the avoidance experiment described above, the mean total errors was 4.68. Equating T in Equation 11 to 4.68 and assuming $N=2$, the resulting c value is .427. This estimate of c will be fixed throughout the remaining discussion and will be used in the calculations of all the following theoretical predictions.

The distribution of T for the two-pattern model, with which we are presently concerned, is given by Equation 10 with $N=2$. The cumulant of the obtained distribution of total errors is given in Figure 3, along with the theoretical prediction.

In deriving further predictions it is useful to have the probabilities, $w_{i,n}$, that the subject is in conditioning state S_i $(i=0,1,2)$ on trial n of the experiment $(n=1, 2, 3, \ldots)$. The result for $w_{0,n}$ is

$$w_{0,n} = (1-c)^{n-1} . \tag{12}$$

For a subject to be in state S_1 on trial n we note that he may remain in state S_0 for k trials $(k=1,2,\ldots,n-1)$ before moving to S_1

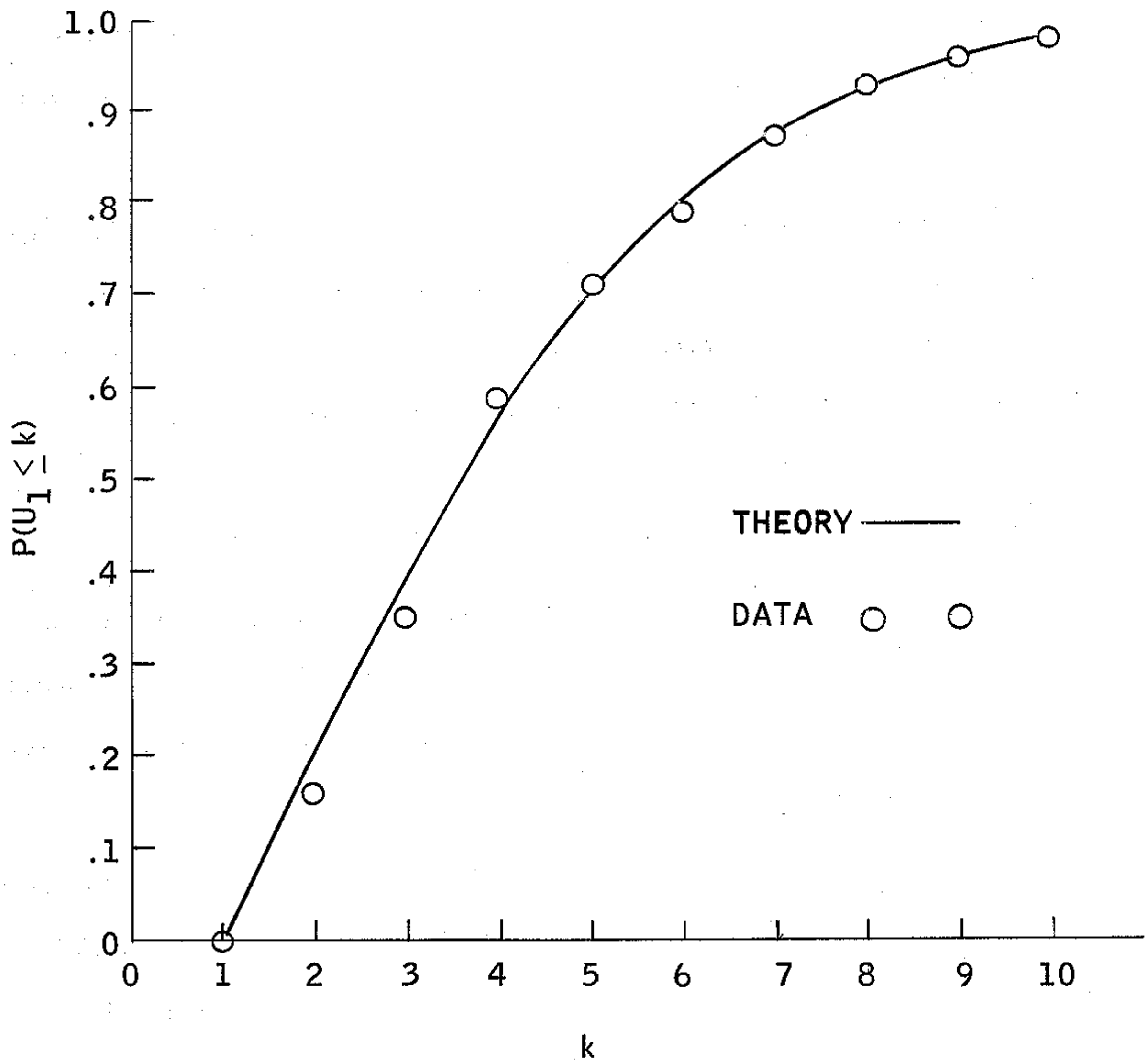

Fig. 3. Cumulant of the Distribution of Total Errors, $P(U_1 \leq k)$.

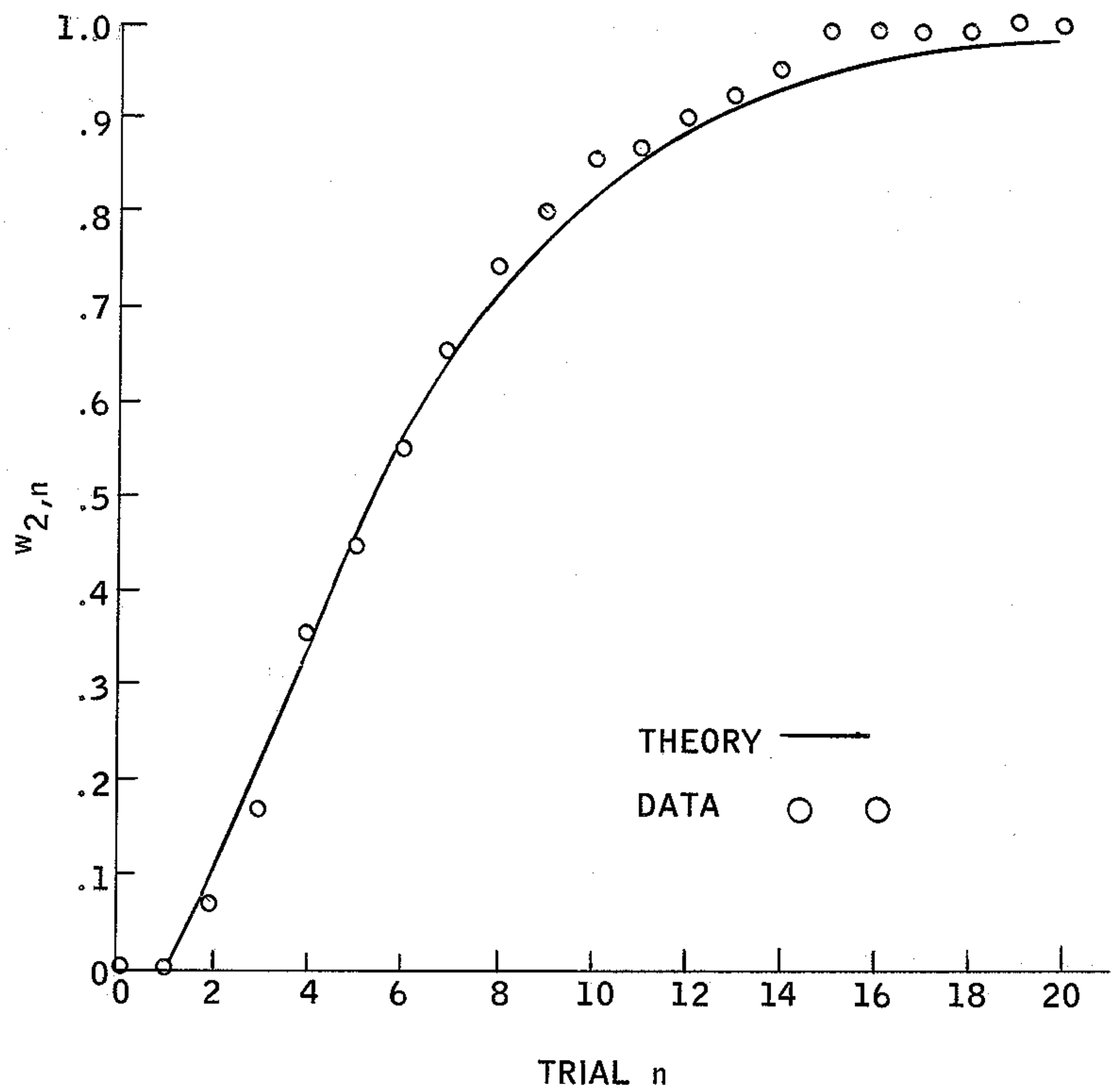

Fig. 4. Proportion of Sequences with no Errors Following Trial $n-1$, ($w_{2,n}$). These Curves also Represent the Cumulant of the Distribution of the Trial Number of the Last Error.

and then have remained in state S_1 for $(n-k-1)$ trials. Thus, the probability of being in state S_1 on trial n will be given by

$$(13) \qquad w_{1,n} = \sum_{k=1}^{n-1} (1-c)^{k-1} c \left(1- \frac{c}{2}\right)^{n-k-1} ,$$

which has the solution

$$(13) \qquad w_{1,n} = 2\left[\left(1- \frac{c}{2}\right)^{n-1} - (1-c)^{n-1}\right] .$$

Having obtained the probabilities of being in conditioning states S_0 and S_1 on trial n, the probability of being in state S_2 may be obtained by subtraction;

$$(14) \qquad \begin{aligned} w_{2,n} &= 1 - w_{0,n} - w_{1,n} , \\ w_{2,n} &= 1 - 2\left[\left(1- \frac{c}{2}\right)^{n-1} + (1-c)^{n-1}\right]. \end{aligned}$$

Once the subject has arrived at conditioning state S_2 there can be no more errors. Hence, $w_{2,n}$ gives the probability of no more errors after trial $n-1$. The observed and theoretical proportions of response sequences having no errors following trial $n-1$ are given in Figure 4. The c value used in the predictions is .427, which was estimated from the mean total errors.

Trial Number of the Last Error

A subject's last error can occur on trial n only if he is in conditioning state S_1, samples the unconditioned pattern, and conditioning of that element is effective. Thus, the probability distribution of the trial number of the last error, $P(L=n)$ for $n=1,2,3, \ldots$, is

$$(15) \qquad P(L=n) = w_{1,n} \cdot \frac{c}{2} = c\,[(1-\frac{c}{2})^{n-1} - (1-c)^{n-1}] ,$$

which has mean and variance

$$(16) \qquad E(L) = \frac{3}{c} , \qquad \mathrm{Var}\ (L) = \frac{(5-3c)}{c^2} .$$

The obtained and predicted mean and standrad deviation of the trial number of the last error are given in Table 3. The cumulant of the distribution of the trial number of the last error is just $w_{2,n}$, the probability of no errors following trial $n-1$, which was given in Figure 4.

The average probability of an error on trial n will be equal to the probability of being in state S_0 on trial n times the probability of an error in S_0 plus the probability of being in state S_1 on trial n times the probability of an error in S_1. Thus,

$$P(A_{2,n}) = 1 \cdot w_{0,n} + \frac{1}{2} \cdot w_{1,n} ,$$

$$(17) \qquad P(A_{2,n}) = (1-c)^{n-1} + \frac{1}{2} \cdot 2[(1-\frac{c}{2})^{n-1} - (1-c)^{n-1}] ,$$

$$P(A_{2,n}) = (1-\frac{c}{2})^{n-1} .$$

The probability of a success on trial n is given by

$$P(A_{1,n}) = 1 - P(A_{2,n}) \, ,$$

(18)

$$P(A_{1,n}) = 1 - (1 - \frac{c}{2})^{n-1} \, .$$

The obtained and predicted mean learning curves are given in Figure 5.

Errors During Various Parts of Learning

The probability of exactly one error before the first success, $P(J_o=1)$, will be equal to the probability that conditioning was effective on the first trial times the probability of sampling the conditioned pattern on the second trial, which is equal to

(19) $$P(J_o=1) = c\,(\frac{1}{2}) \, .$$

Two errors before the first success can occur if conditioning is not effective until trial 2 and then the conditioned pattern is sampled on trial 3, or if conditioning is effective on trial 1, the unconditioned pattern is sampled on trial 2, and then either (a) the second pattern becomes conditioned on trial 2, or (b) ineffective conditioning on trial 2 and the single conditioned pattern is sampled on trial 3. Thus the probability of two errors before the first success will be given by

$$P(J_o=2) = (1-c)c\frac{1}{2} + c\frac{1}{2}c + c\frac{1}{2}(1-c)\frac{1}{2}$$

(20)

$$P(J_o=2) = \frac{1}{2}c\,[1 + \frac{1}{2}\,(1-c)] \, .$$

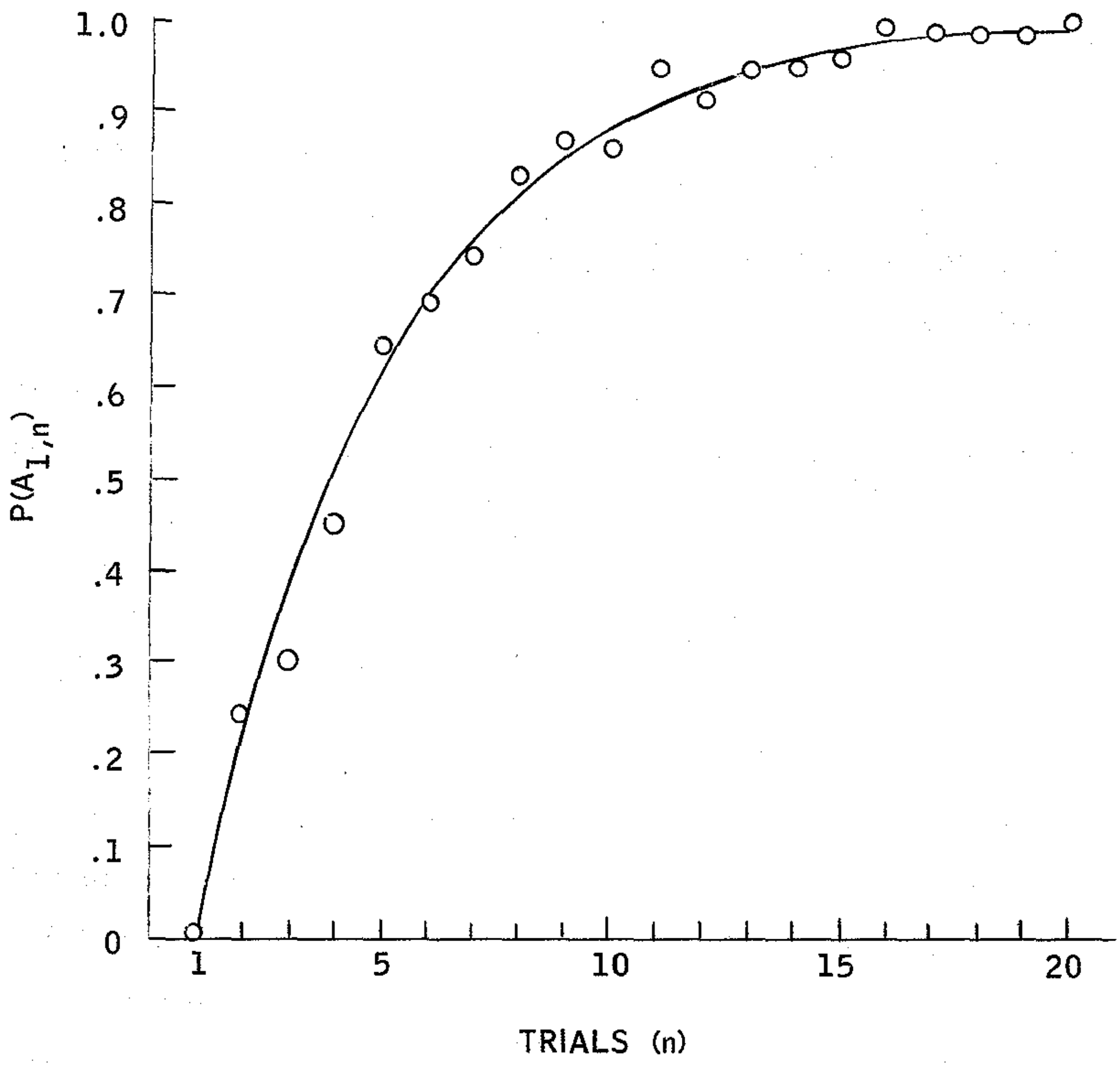

Fig. 5. Probability of an Avoidance Response, $P(A_{1,n})$, on Successive Trials of the Experiment.

By the same reasoning, it follows that the probability of obtaining three errors before the first success is

(21) $$P(J_0=3) = \frac{1}{2}c\ (1-c)\ [1 + \frac{1}{2} + \frac{1}{4}\ (1-c)]\ .$$

In general, it can be shown by induction that the probability that J_0 takes on an arbitrary value, k, will be given by

(22) $$P(J_0=k) = \begin{cases} \frac{1}{2}\ c & \text{for } k = 1 \\ c(1-c)^{k-2}\ [1 - (1+c)(\frac{1}{2})^k] & \text{for } k \geq 2 \end{cases}$$

which has a mean equal to

(23) $$E(J_0) = \frac{1}{c} + \frac{1}{1+c}\ .$$

The cumulants of the obtained and predicted distributions of the number of errors before the first success are given in Figure 6.

A sequence with no reversals can be defined as a response sequence in which no errors occur after the first success. Sequences with no reversals can occur only if there have been at least two initial errors, since conditioning occurs only on trials on which an error was made and there are two stimulus patterns which must become conditioned. A sequence with two errors followed by all successes can occur by conditioning being effective on the first trial, sampling the unconditioned pattern on the second trial, and then having conditioning again effective. Thus, the probability of obtaining a non-reversal sequence with two errors, $P(NR=2)$, is

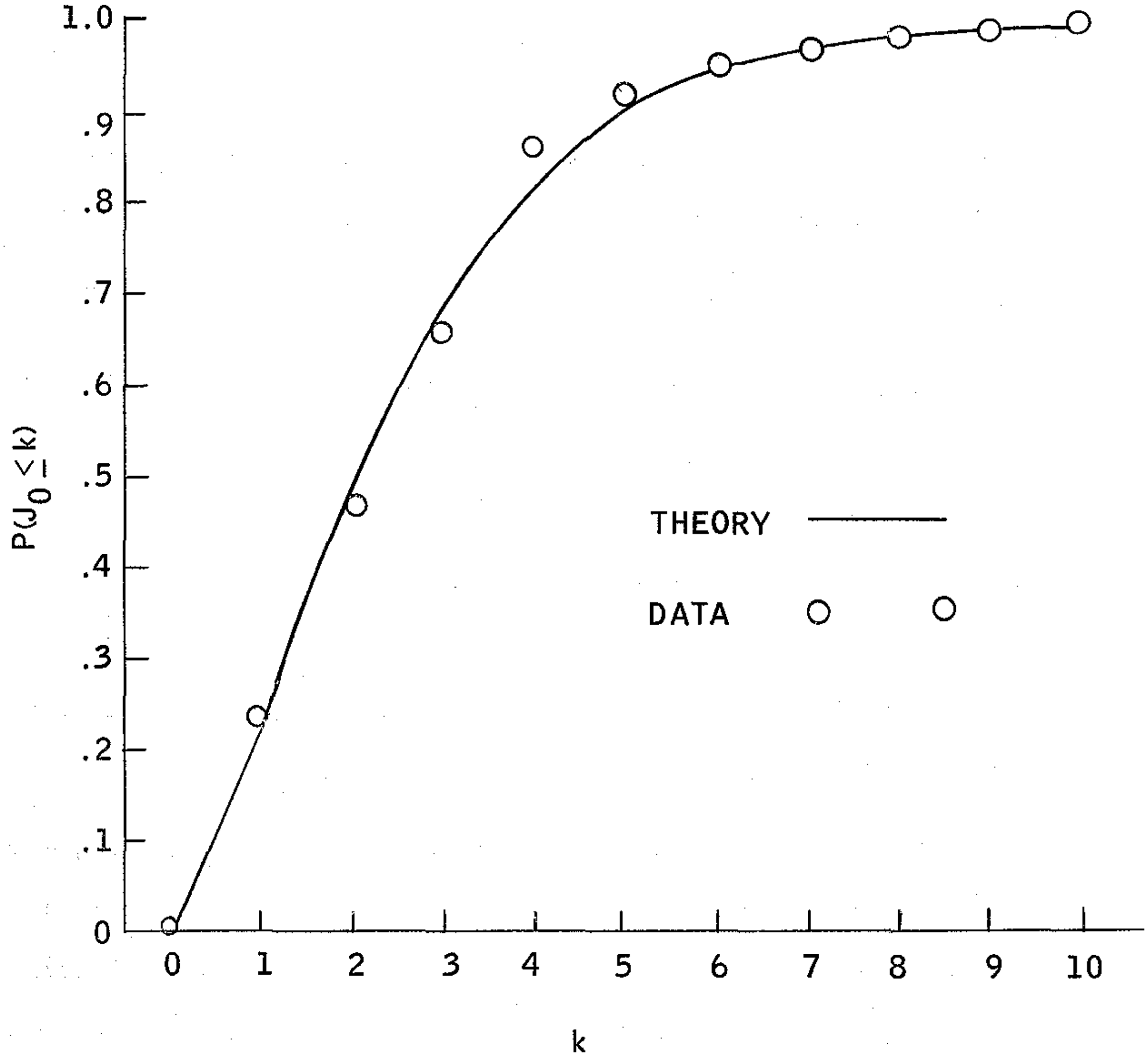

Fig. 6. Cumulant of the Distribution of Errors Before the First Success, $P(J_0 \leq k)$.

(24) $$P(NR=2) = \frac{c^2}{2} .$$

A non-reversal sequence with three errors can come about by ineffective conditioning on the first trial followed by effective conditioning on the next two trials, or by effective conditioning on the first trial, sampling the unconditioned pattern on the second trial with conditioning ineffective, and then sampling it again with conditioning effective. Therefore, the probability of a non-reversal sequence with three errors is

(25) $$P(NR=3) = c^2(1-c)(\frac{1}{2} + \frac{1}{4}) .$$

Similarly, the probability of obtaining a non-reversal sequence with four errors will be

(26) $$P(NR=4) = c^2(1-c)^2 \ (\frac{1}{2} + \frac{1}{4} + \frac{1}{8}) .$$

In general, it can be shown by induction that the probability of obtaining k errors (k=2,3,4, ...) in a non-reversal sequence will be given by

(27) $$P(NR=k) \doteq c^2(1-c)^{k-2} \ [1 - (\frac{1}{2})^{k-1}] .$$

The probability of obtaining a sequence without any reversals, P(NR), will just be

$$(28) \quad P(NR) = \sum_{k=2}^{\infty} P(NR=k) = c^2 \left[\sum_{k=2}^{\infty} (1-c)^{k-2} - \frac{1}{2} \sum_{k=2}^{\infty} \left(\frac{1-c}{2}\right)^{k-2} \right],$$

which has the solution

$$(29) \qquad P(NR) = \frac{c}{1+c} .$$

Obtained and predicted values for P(NR) are given in Table 4.

Errors between adjacent successes can only occur in state S_1. If we let g_k represent the probability that the k-th success occurred in state S_1, then g_1 is

$$(30) \qquad g_1 = 1 - P(NR) = \left(\frac{1}{1+c}\right) .$$

The general equation for g_k will be

$$(31) \qquad g_k = g_{k-1} \sum_{i=0}^{\infty} \left(\frac{1-c}{2}\right)^i \frac{1}{2} = g_{k-1} \left(\frac{1}{1+c}\right) = \left(\frac{1}{1+c}\right)^k .$$

Letting J_k be the number of errors between the k-th and the (k+1)-st successes, the probability distribution of J_k is

$$(32) \qquad P(J_k=i) = \begin{cases} 1 - \frac{1}{2}g_k & \text{for } i = 0 \\ \frac{1}{4}g_k(1+c)[\frac{1}{2}(1-c)]^{i-1} & \text{for } i \geq 1 \end{cases},$$

and the expectation of J_k will be

$$(33) \qquad E(J_k) = (\frac{g_k}{1+c}) = (\frac{1}{1+c})^{k+1} \qquad \text{for } k \geq 1 .$$

The obtained and expected mean number of errors between the k-th and (k+1)-st successes are plotted in Figure 7 for k going from 0 to 9.

The cumulative errors before the k-th success, F_k, can be obtained by adding the values of J_i from i=0 to i=k-1.

The expectation of F_k will be

$$(34) \qquad E(F_k) = E(J_o) + \sum_{i=1}^{k-1} E(J_i) = \frac{2}{c} - \frac{1}{c} [\frac{1}{1+c}]^k .$$

The limit of $E(F_k)$ as k approaches infinity is the expected mean total errors, (2/c), as it should be.

Let p_k be the probability of no errors following the k-th success. Successes can occur only when the subject is in conditioning states S_1 or S_2. If the k-th success occurs in S_1 there will be at least one more error. Therefore, to be followed by no more errors, the k-th success must occur in state S_2, which it can with probability

$$(35) \qquad p_k = 1 - g_k = 1 - (\frac{1}{1+c})^k , \qquad \text{for } k = 1,2,3, \ldots .$$

Define a random variable Z taking the values 0, 1, 2, 3, ... , which represents the number of successes before the subject's last error. The probability distribution of Z can be determined from successive differences of the p_k values of Equation 35, viz.,

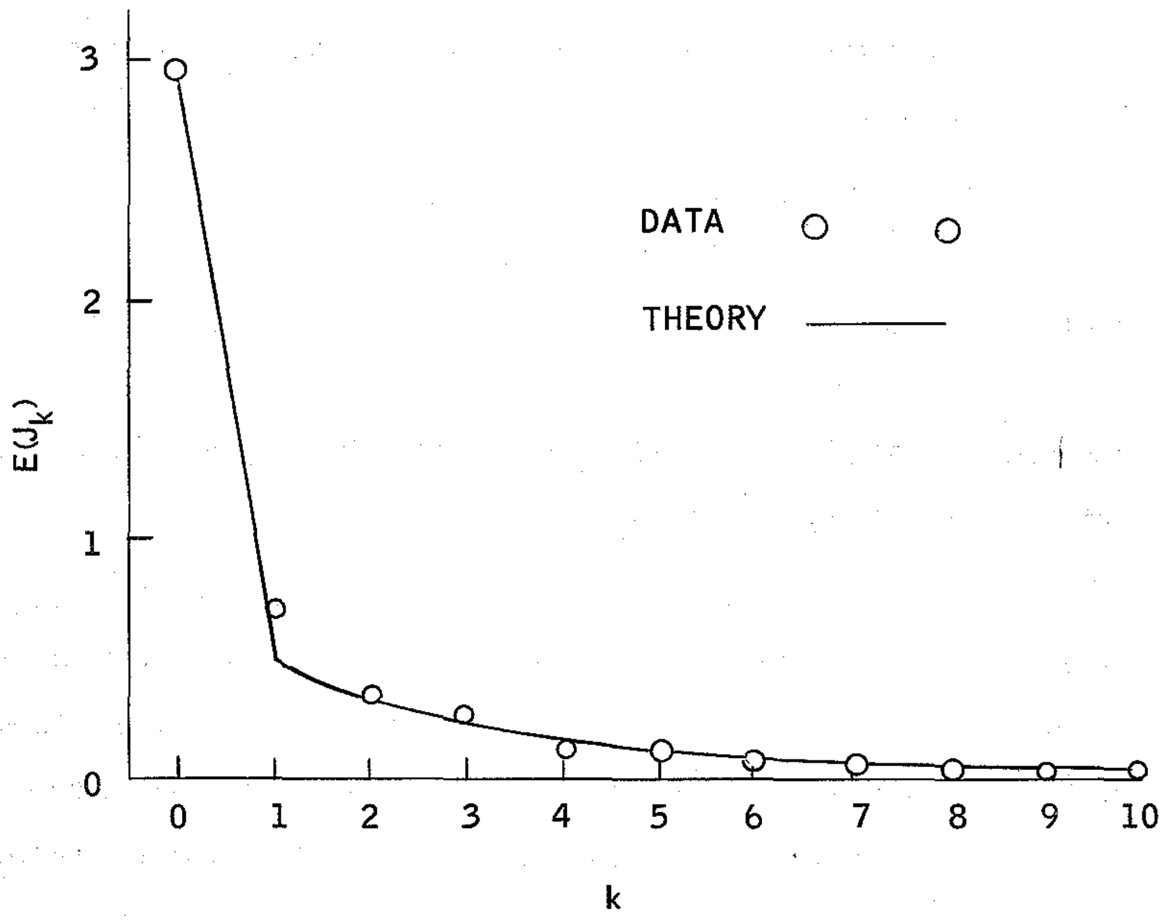

Fig. 7. Mean Number of Errors Between the k^{th} and $(k+1)^{st}$ Successes, $E(J_k)$.

(36) $$P(Z=k) = p_{k+1} - p_k = c[\frac{1}{1+c}]^{k+1} , \qquad \text{for } k=0,1,2, \ldots .$$

This distribution of Z has mean and variance equal to

(37) $$E(Z) = \frac{1}{c} , \qquad Var(Z) = \frac{1+c}{c^2} .$$

The obtained and predicted distributions of Z are given in Figure 8.

Sequential Statistics

In deriving sequential statistics it is useful to define a sequence of response random variables, X_n, which take on the value 1 if an error occurred on trial n or the value 0 if a success occurred on trial n. From the axioms, the conditional probabilities of an error given state S_i are

(38) $$P(X_n=1|S_{0,n}) = 1 , \quad P(X_n=1|S_{1,n}) = \frac{1}{2} , \quad P(X_n=1|S_{2,n}) = 0 .$$

Using the notation of Bush (1959), a j-tuple of errors will be defined as u_j, where

(39) $$u_{j,n} = X_n \cdot X_{n+1} \cdot \ldots \cdot X_{n+j-1} \qquad \text{for } j=1,2, \ldots$$

and

(40) $$E(u_{j,n}) = w_{0,n} \left\{(1-c)^{j-1} + \sum_{k=1}^{j-1} (1-c)^{k-1} \frac{1}{2}c\left[\frac{1}{2}(1-c)\right]^{j-1-k}\right\} + w_{1,n}[(\tfrac{1}{2})^{j} (1-c)^{j-1}] .$$

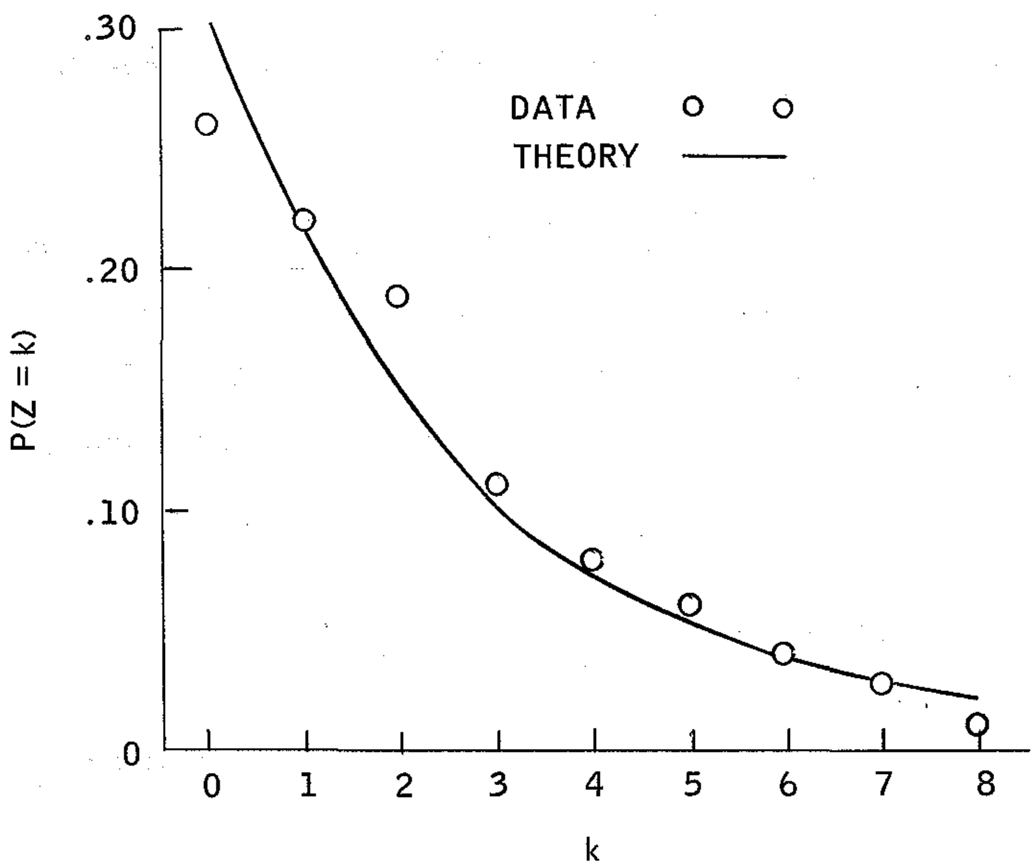

Fig. 8. Probability of k Successes Before the Last Error, P(Z = k).

The expected number of j-tuples of errors will have the solution

$$(41) \qquad E(u_j) = E\left[\sum_{n=1}^{\infty} u_{j,n}\right] = \frac{(1-c)^{j-2}}{c}\left[1 + (1-2c)\left(\frac{1}{2}\right)^{j-1}\right] \quad \text{for } j \geq 2 .$$

The value of u_1 will be given by mean total errors, which has expectation $2/c$. Bush (1959) has shown that predictions about runs of errors can be obtained once the expected u_j are known. Defining total error runs as R, the expectation of R is

$$(42) \qquad E(R) = E(u_1) - E(u_2) = 1 + \left(\frac{1}{2c}\right) .$$

Letting r_j be the number of error runs of length j, for $j=1,2,3, \ldots$, the expectation of r_j is

$$(43) \qquad E(r_j) = E(u_j) - 2E(u_{j+1}) + E(u_{j+2}) ,$$

$$E(r_j) = \begin{cases} \frac{1}{4}c\,(1+c+2c^2) , & \text{for } j = 1 \\ c(1-c)^{j-2} + \left[\dfrac{(1-c)(1+c)^2\left(\frac{1-c}{2}\right)^{j-2}}{8c}\right] , & \text{for } j \geq 2 . \end{cases}$$

The number of alternations of errors and successes, A, over the entire learning sequence will be twice the number of error runs minus one. Thus, the expectation of A will be

$$(44) \qquad E(A) = 2 \cdot E(R) - 1 = 1 + \left(\frac{1}{c}\right) .$$

The obtained and predicted values for runs of errors are given in Table 4.

Another useful summary of sequential characteristics in the response data is the extent to which an error on trial n tends to be followed by an error k trials later, irrespective of the intervening responses. Define the autocorrelation, $c_{k,n}$, as the product $X_n \cdot X_{n+K}$, which will have the value 1 if errors occurred on both trials n and $n+k$ and the value 0 otherwise. The expectation of $c_{k,n}$ is

$$(45)\quad E(c_{k,n}) = [w_{0,n}(w_{0,k+1} + \frac{1}{2} w_{1,k+1})] + [w_{1,n} \frac{1}{4} (1- \frac{1}{2}c)^{k-1} (1-c)] .$$

The expectation of $c_{k,n}$ over all trials will be

$$(46)\quad c_k = E\left[\sum_{n=1}^{\infty} c_{k,n}\right] = \frac{(3-2c)}{2c} (1- \frac{1}{2}c)^{k-1} ,$$

for $k=1,2,3, \ldots$. The obtained and predicted autocorrelations of errors for the first few values of k are given in Table 3.

Table 3

Observed and Predicted Values for Various Response Measures

Response Measure	Observed	Predicted
Total errors		
$E(T)$	4.68	----
$\sigma(T)$	2.34	2.48
Trial number of last error		
$E(L)$	6.56	7.02
$\sigma(L)$	3.40	4.52
Errors before the first success		
$E(J_0)$	2.96	3.03
$\sigma(J_0)$	1.83	2.14
Probability of no reversals		
$P(NR)$	.26	.30
Mean number of runs of errors		
R	2.18	2.18
Runs of errors, r_j, of length j.		
r_1	1.05	1.05
r_2	.46	.51
r_3	.30	.27
Autocorrelation of errors, c_k, k trials apart		
c_1	2.50	2.62
c_2	2.06	2.06
c_3	1.47	1.61
c_4	1.18	1.27
Mean length of runs of successes in state S_1		
$E(\bar{h})$	1.77	2.00
$\sigma(\bar{h})$	1.14	1.41

THE SOLOMON AND WYNNE DATA

Having achieved a high degree of success in applying the two-pattern Markov model to avoidance conditioning in rats, the model will now be applied to data from an experiment by Solomon and Wynne (1953) where 30 dogs learned an avoidance response. This application of the model is of great interest since Bush and Mosteller (1959) have already fitted eight models to this data, with varying degrees of success. It should be noted beforehand that the procedure used by Solomon and Wynne does not provide an optimum test of the two-pattern model. For example, the use of the shuttle response, which requires the subject to jump to a place where he has been previously shocked, and the spacing of trials would tend to complicate the stimulus situation. In spite of these drawbacks in the Solomon and Wynne experiment, it is felt that important information on the generality of the Markov model could be gained by applying it to the dog data. It is fortunate that Bush and Mosteller (1955) have published response sequences of Monte Carlo, stat-dogs having the properties of a two-operator linear model, with parameters estimated from the real dog data. Thus, the real dog data can be compared to both the predictions from the Markov model and the values from the linear stat-dogs.

Bernoulli Properties

The Markov model predicts that the probability of a success (an avoidance response) between the first success and the last error (failure to avoid) should be a constant, 1/2. According to the linear model, the probability of a success should not be stationary, but should increase during these trials. The probability of a success on trials between the first success and the last error (excluding the last error) for the real and linear stat-dogs are given in Figure 9. It can be seen that the real dogs are approximating a constant value near .57, while the linear stat-dogs show a systematic increase over trials. Thus, the real dogs exhibit the Markov property of stationarity, but slightly above .50.

According to the Markov model, the conditional probability of a success following a success on these Bernoulli trials should be a constant, .500. This conditional probability for the real dogs was .492, but was .626 for the linear stat-dogs.

The response sequences between the first success and the last error for the real and stat-dogs were divided into blocks of three trials. According to the Markov model the number of errors in each block should be distributed binomially, with $p = \frac{1}{2}$. Goodness of fit test of the real data to the theoretical predictions yielded a Chi Square of .89 with 3 degrees of freedom, which is not significantly different from what would be expected on the basis of chance. On the other hand, the comparison of the linear stat-dog data to the binomial predictions yielded a Chi Square of 16.08 with 3 degrees of freedom, which is significant at beyond the .05 level. Thus, a binomial response sequence characterizes the real data, but does not characterize the linear stat-dog data.

Further, a test of the two-pattern prediction that the response sequence from the first success to the last error should be a zero-order

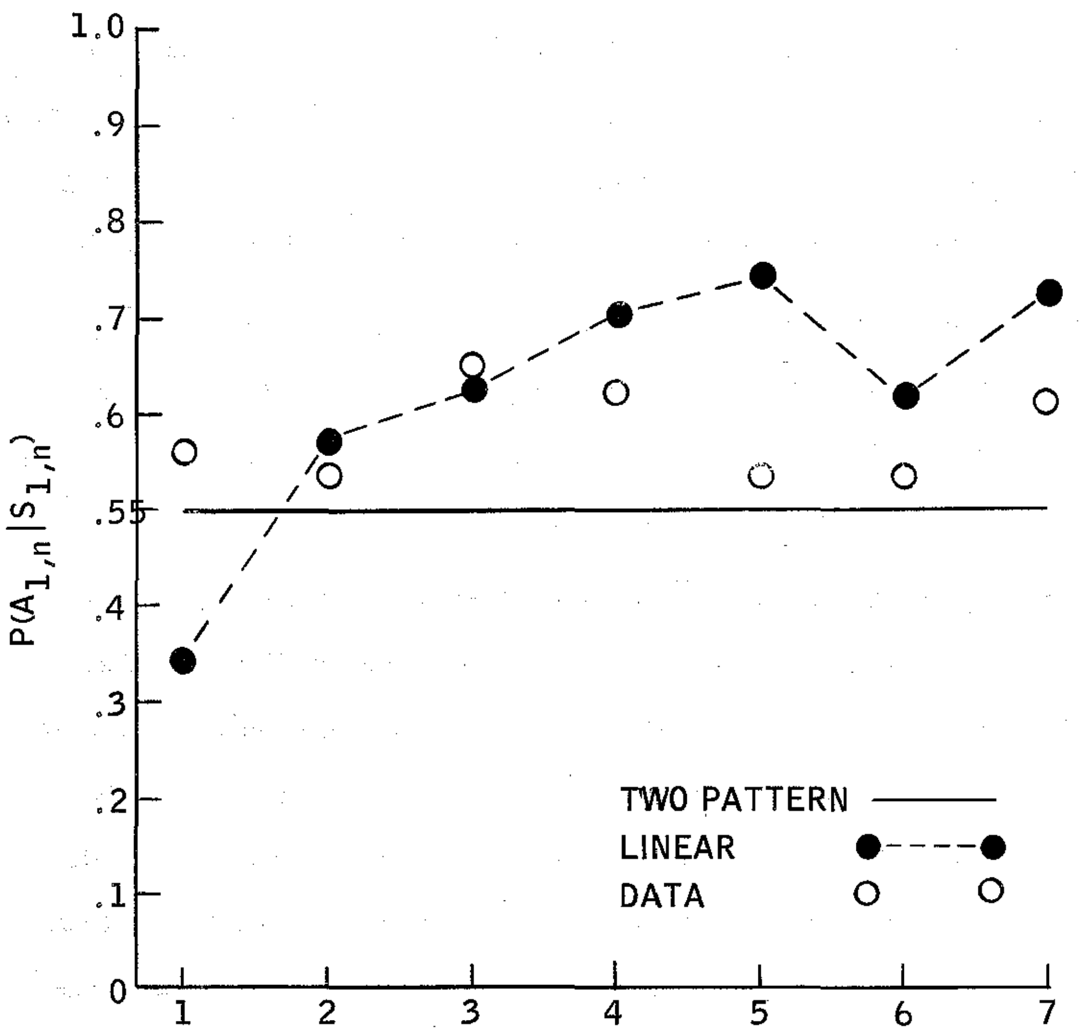

Fig. 9. Probability of a Success on Trials Between the First Success and the Last Error for the Solomon and Wynne Dogs and the Bush and Mosteller Linear Stat-Dogs.

Markov process yielded a nonsignificant Chi Square of 1.79 with one degree of freedom. Thus, we cannot reject the hypothesis that the response sequence on these trials is a zero-order Markov process.

Other Predictions

Table 4 presents a comparison of the two-pattern Markov predictions and the data from the real and linear stat-dogs on various response measures. Mean total errors, $E(T)$, was used to estimate the conditioning parameter, c, which is .257. It can be seen that the Markov model is predicting the real data as well or better than the two-operator linear model on some of the response measure. The Markov model's glaring inadequacy is the systematic overestimation of the standard deviations.

Small standard deviations would result if the stimulus situation consisted of more than two stimulus patterns. The obtained mean and variance of the total errors can be set equal to the general theorical equations for the mean and variance of the total errors (Equation 11), and the two equations can be solved simultaneously for the conditioning parameter, c, and the number of stimulus patterns, N. For the Solomon & Wynne data, this estimate of c is .538, and the estimated number of stimulus patterns, N, is 4.2. Thus, it is possible that a four-pattern model would describe the data better than the two-pattern model. The obtained probability distribution of total errors is given in Figure 10, along with the predictions from the two-pattern and the four-pattern models. It can be seen that the four-pattern model describes the data better than does the two-pattern model.

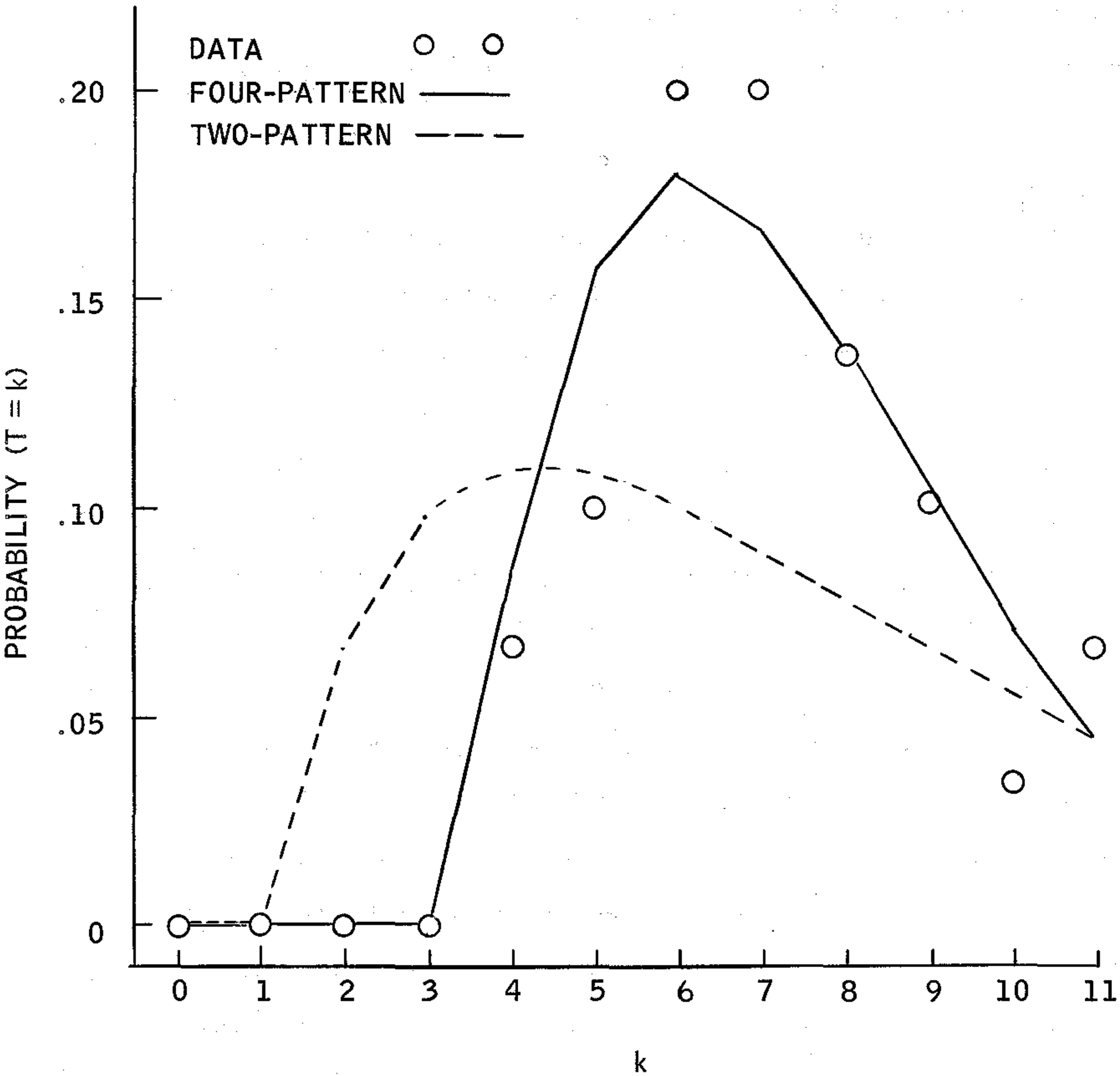

Fig. 10. Probability Distribution of Total Errors for the Solomon and Wynne Dogs and the Predictions from the Two-Pattern and Four-Pattern Markov Models.

The four-pattern predictions for the mean and standard deviations of the total errors and the trial number of the last error are given in Table 4. Because of the difficult mathematics involved in the four-pattern model, the predictions for the other statistics are not yet available. It can be seen that the four-pattern model is predicting the standard deviations better than the linear or two-pattern models, but the prediction for the mean of the trial of the last error is off.

Since the Solomon & Wynne situation (four-pattern) involved the bidirectional, shuttle response, and the rat experiment (two-pattern) involved a undirectional response, it is possible that each side of the shuttle box can be represented by two stimulus elements. Assuming that the learning on each side of the box is independent, the two-pattern model could be applied independently to the odd and even trials. This hypothesis needs to be tested under conditions which are optimum for the two-pattern model, e.g. uniform spacing of trials.

DISCUSSION

A two-pattern stimulus sampling model, which can be represented as a three state, absorbing Markov chain, has been presented to account for simple conditioning. A large number of predictions about sequences of response random variables were derived in closed form and were compared to actual data obtained in an avoidance conditioning experiment with rats. The theoretical predictions fit the rat data extremely well. In particular, the predicted Bernoulli properties of the response sequences between the first success to the last error were upheld suprisingly well by the data. This characteristic of the

Table 4

A Comparison of Avoidance Conditioning Data of Real Dogs with Linear Stat-dogs and Markov Model Predictions

Response Measure		4-Pattern	2-Pattern	Data	Linear
Total errors,	$E(T)$	7.43	----	7.80	7.60
	$\sigma(T)$	2.53	4.75	2.52	2.27
Errors before k-th success,					
	$E(F_1)$	----	4.68	4.50	4.13
	$\sigma(F_1)$	----	3.66	2.25	2.08
	$E(F_2)$	----	5.32	5.47	5.20
	$E(F_3)$	----	5.81	6.10	5.94
Trial of last error	$E(L)$	15.47	11.70	12.33	13.53
	$\sigma(L)$	4.62	8.02	4.36	4.78
Total runs of errors,	R	----	2.96	3.30	3.47
Runs of errors,	r_1	----	1.37	1.86	1.87
	r_2	----	.63	.40	.57
	r_3	----	.33	.27	.27
Autocorrelations,	c_1	----	4.86	4.50	4.11
	c_2	----	4.25	4.36	3.53
	c_3	----	3.70	3.40	3.14
Proportion of successes on two trials prior to the last error			.50	.52	.78
Probability of no errors after 1-st success			.20	.10	.03

data is sufficient to question the adequacy of any linear model as a description of the present rat data. It should be noted also, that no estimated parameters entered into the predictions about the response sequences during trials between the first success and the last error, since the Bernoulli properties of the model are parameter-free.

The predictions about various response measures over the entire learning sequence involved only a single parameter, c. Although c was estimated from an arbitrary response measure, mean total errors, the predictions for the numerous other response measures fit the data very well. It should be noted that entire distributions of response variables were fitted to the data, and not just mean values.

As a test of its generality, the two-pattern Markov model was applied to the data of Solomon & Wynne (1953) on avoidance conditioning in dogs. Again the Bernoulli properties of the model were upheld by the dog data, but the obtained probability of an avoidance response on the Bernoulli trials was slightly higher than the predicted value of 1/2. In general, the two-pattern model fit the dog data reasonably well in spite of the fact that the experimental procedure did not afford an optimum test of the model. The variances predicted by the two-pattern model were too large and indicated that some type of four-pattern Markov model might fit the data better. The strongest evidence for the two-pattern Markov model as an interpretation of the dog data is the fact that the response sequences between the first success and the last error are independent Bernoulli trials. In view of this fact, it is surprising, and somwhat paradoxical, that the two-operator linear model predicts means and variances of other response measures as well as it does. The Bernoulli predictions certainly do not follow from the linear model. The analyses of the Solomon & Wynne data indicate (a) that the two-pattern Markov model cannot be excluded as a possible interpretation of these data, (b) that a four-pattern process might be involved here, and (c) that the two-operator linear model does not give a complete description of the data, even though it closely approximates the data in many respects.

The two-pattern Markov model has been presented as applying to all simple conditioning situations. However, to date it has been applied only to avoidance learning data. The model predicts that learning occurs only on trials on which an error is made, and in fact, in avoidance conditioning an experimenter-controlled reinforcement occurs only on these error trials. Therefore, it is of interest whether this assumption of learning only on error trials will be upheld by, say, classical conditioning or T-maze reversal data, where an experimenter-controlled reinforcement occurs on both error and success trials.

APPENDIX

Table 5

Response Sequences for Individual Rats

Code: The columns represent consecutive runs of failures to avoid (errors) and runs of avoidance responses (successes). The columns labeled "1" designate runs of errors. The columns labeled "0" designate runs of successes. The entries of the table give the length of the runs in number of trials. The final criterion run of 20 successes has not been indicated, but it follows the last entry for each rat.

Original Learning

Rat	1	0	1	0	1	0	1
1	4	1	1	1	1		
2	4	1	3				
3	3	4	1				
4	4	5	1				
5	6	2	2				
6	2	1	1	1	1		
7	4						
8	2	1	1				
9	1	1	2				
10	4	2	2	1	1		
11	3	1	2				
12	3	1	1				
13	5	1	1				
14	4						
15	4						
16	3	2	1				
17	2						
18	1	1	1	2	1	3	1
19	2	1	4	1	1		
20	5						
21	2	1	4	2	1		
22	2	5	1				
23	5	1	3	2	2		
24	2	1	1				
25	4	4	2				
26	3						
27	1	1	3				
28	2	1	1				
29	3	1	1				
30	5	2	1				
31	5						
32	3	1	1	1	2		
33	4	1	1				
34	1	1	1				
35	3	1	1				
36	1	2	1				
37	1	1	1	1	3		
38	1	3	2				
39	4						
40	5						
41	1	2	1				
42	2						
43	1	1	3	4	1	1	1
44	3	1	2	2	1	4	2
45	6	1	1	5	2		
46	4						
47	2	3	2	5	3		
48	8						
49	4						
50	3	2	1				

Reversal Learning

Rat	1	0	1	0	1	0	1	0	1
1	6	3	3						
2	2								
3	2	1	1						
4	3	1	1	1	1	1	1		
5	2	1	3						
6	1	1	1						
7	3								
8	1	1	1						
9	4	1	2	3	1				
10	3								
11	1	2	1						
12	1	1	1	4	1				
13	4								
14	3	1	2						
15	1	1	1	1	1	1	1		
16	2	1	1						
17	1	1	1						
18	4	2	1	1	1				
19	7								
20	9	1	1	1	1				
21	1	1	2	1	1	3	2	2	1
22	2								
23	2	1	1	2	1				
24	4	3	1	1	1	1	2		
25	3	2	1	2	1	2	1	1	1
26	3								
27	2	3	1						
28	1	1	1	1	1				
29	2								
30	1	3	1	1	1				
31	1	1	1	1	2				
32	7	3	2						
33	3	5	1						
34	4								
35	1	1	3						
36	2	3	1						
37	2								
38	1	2	1						
39	4	2	4						
40	10	1	2	3	3				
41	3	2	3	1	1	1	1	2	1
42	1	1	1	3	2				
43	4								
44	1	3	1	1	1				
45	3								
46	2								
47	2	1	1	4	2				
48	1	1	1						
49	2	1	2	1	1				
50	4	2	1						

REFERENCES

Bower, G. H. Properties of the one-element model as applied to paired associate learning. Technical Report No. 31, 1960, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Bush, R. R. Sequential properties of linear models. In Bush, R. R. and Estes, W. K. (Ed.): Studies in mathematical learning theory. Stanford University Press, 1959.

Bush, R. R. and Mosteller, F. Stochastic models for learning. New York: Wiley, 1955.

Bush, R. R. and Mosteller, F. A comparison of eight models. In Bush, R. R. and Estes, W. K. (Ed.): Studies in mathematical learning theory. Stanford University Press, 1959.

Estes, W. K. Toward a statistical theory of learning. Psychol. Rev., 1950, 57, 94-104.

Estes, W. K. Learning theory and the new mental chemistry. Psychol. Rev., 1960, 67, 207-223.

Hull, C. L. Principles of Behavior. Yale University Press, 1943.

Solomon, R. L. and Wynne, L. C. Traumatic avoidance learning: Acquisition in normal dogs. Psychol. Monogr., 1953, 67, No. 4, (Whole No. 354).

Suppes, P. and Atkinson, R. C. Markov learning models for multi-person interactions. Stanford University Press, 1960.