PROPERTIES OF THE ONE-ELEMENT MODEL AS APPLIED TO PAIRED ASSOCIATE LEARNING

by

GORDON H. BOWER

TECHNICAL REPORT NO. 31

June 29, 1960

PREPARED UNDER CONTRACT Nonr 225(17)

(NR 171-034)

FOR

OFFICE OF NAVAL RESEARCH

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

PROPERTIES OF THE ONE-ELEMENT MODEL AS APPLIED
TO PAIRED ASSOCIATE LEARNING

by

Gordon H. Bower

This paper presents the results of exploring the consequences of an elementary theory of learning. Recently the model has been successfully employed in the analysis of the data from several experiments carried out by the writer on paired associate learning of human subjects. Because of the simplicity of the model, a large number of consequences (theorems) can be easily derived and compared with the data. To date these predictions have fared very well in accounting for data from the writer's experiments on paired associates learning. In addition, it has been found that the structure of the model is such that reasonably natural hypotheses may be stated relating parameters of the model to experimental variables.

The adequacy of the model in accounting for the data of a number of preliminary experiments is sufficient justification for exploring in detail the consequences of the model. In purpose, then, this report seeks to do for the one-element learning model what the Bush and Sternberg paper [1] did for the single operator linear model; viz., derive an extensive number of predictions from the axioms. The derivations of results from the one-element model are generally easier than analogous derivations from the linear model and the resulting expressions are usually simpler in form.

The learning model to be discussed here is one special case of a large class of small element stimulus sampling models originally explored extensively by Suppes and Atkinson [2] and Estes [3]. For this class of models, it is assumed that the stimulus situation for the learning subject may be represented by a small number of separate aspects or elements and that each element is conditioned to one or another mutually exclusive responses available to the subject. The subject is assumed to sample at random one of the stimulus elements on each trial and makes the response to which that sampled element is conditioned. At the termination of the trial one or another response is reinforced and with probability θ the sampled element becomes conditioned to the reinforced response, if it was not already conditioned to that response. One general consequence of these small element models is that response probability can change only by discrete steps, the number of steps determined by the number of stimulus elements. In the early chapters of their book, Suppes and Atkinson show general methods of derivations for models constructed on the assumption of one element, of two elements, and, in general, N elements; they further show applications of these models to the results of experiments on multi-person interaction situations. The writer in carrying out his analysis of the present model has borrowed freely from the general derivational techniques displayed in the Suppes and Atkinson book.

The basic notion of the present model is that each stimulus item in the list of paired associates may be represented by exactly one stimulus element within the model and that the correct response to that item becomes associated in all-or-none fashion. Considering only a single item it can be in either of two "states" on each trial: conditioned to the correct response or not conditioned to the correct response. The effect of a reinforced trial (i.e., evoking the correct response in the presence of the stimulus item) is to provide an opportunity for the item to become conditioned. The single parameter of the model is θ, the probability that an unconditioned item will become conditioned as the result of a reinforced trial. All items begin in the unconditioned state and the effect of continued reinforced training is to increase the probability that the item has become conditioned.

If the item has become conditioned, then continued reinforcements of the same correct response will ensure that the item remains conditioned. The probability of the correct response when the item is conditioned is unity. The probability of the correct response when the item is not conditioned depends upon the exact experimental procedure used. In experiments by the writer, the subjects were told the N responses (integers $1, 2, \ldots, N$) available to them and were told to respond on every trial regardless of whether they knew the correct number. If the N numbers occur equally often as the to-be-learned responses to the items, then the probability that the subject will guess correctly on an unlearned item is $1/N$; correspondingly, his probability of guessing incorrectly is $1 - 1/N$. Our discussion of the one-element model is oriented specifically towards such an experimental procedure.

A frequently used alternative procedure in paired-associate learning experiments requires the subject to learn the response alternatives (e.g. words or nonsense syllables) as well as the S-R associations. In addition, the exposure-time to each item is usually controlled by the experimenter and the trial-terminating event (i.e., correct response in paired associate, next item in the list in rote serial learning) follows inexorably regardless of whether the subject makes a relevant response. The result of this procedure is that early in training the subject's protocol includes a number of failures to respond and irrelevant responses which are not members of the set of available alternative responses. Detailed theoretical analysis of this type of data will probably prove quite difficult and no such attempt is made here.

Because of the way the model is formulated, there is a partial determinism between the response sequence and the sequence of conditioning states. Specifically, if the subject responds incorrectly to a given item on trial n, then we know that that item was not in the "conditioned" state on trial n. This feature is very helpful in deriving a number of the theorems about errors. If the subject responds correctly, however, then we cannot uniquely specify his state of conditioning since he may have guessed correctly. Thus, it is not a consequence of the model that the subject's first correct response will be followed with probability 1 by correct responses on subsequent trials.

With these preliminary remarks, we now turn to the statement of the axioms and derivations of results. The first section is concerned with the analysis of the sequence of correct and incorrect responses to a given item over successive trials of the experiment. The concluding section is concerned with the derivation of statistical features within the total list of items considering any single trial of the experiment.

II. Axioms and theorems about total errors.

Axioms

1. Each item may be represented by a single stimulus element which is sampled on every trial.
2. This element is in either of two conditioning states: C_1 (conditioned to correct responses) or C_0 (not conditioned).
3. On each reinforced trial, the probability of a transition from C_0 to C_1 is a constant, θ; the probability of a transition from C_1 to C_1 is 1.
4. If the element is in state C_1 then the probability of a correct response is 1; if the element is in state C_0, then the probability of a correct response is $1/N$, where N is the number of response alternatives.
5. The probability θ is independent of the trial number and the outcomes of preceding trials.

The trial sequence of states of conditioning forms a Markov chain, with C_1 being an absorbing state. The transition probabilities are given in the following matrix:

$$\text{(1)} \qquad P = \begin{array}{c|cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 0 & \theta & 1-\theta \end{array}$$

It is easy to show that the n^{th} power of the transition matrix is

$$\text{(2)} \qquad P^n = \begin{array}{c|cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 0 & 1-(1-\theta)^n & (1-\theta)^n \end{array}$$

We explicitly assume that all items start out in state 0 (i.e., are not conditioned initially). Thus, starting out in state 0, the probability of still being in state 0 after n reinforced trial is $(1-\theta)^n$, which approaches zero as n becomes large. Thus, for $0 < \theta < 1$, with probability 1 the process will eventually end up in conditioning state C_1 (i.e., will become conditioned).

What is the average number of trials the process will be in state 0 before conditioning occurs? To answer this question, define a random variable f_n to be 1 if the process is in state C_0 at the beginning of trial n, and f_n is 0 otherwise. Then the average number of times, $\bar{n}_0$, the process is in state 0 is

(3) $$\bar{n}_0 = E[\sum_{n=1}^{\infty} f_n] = \sum_{n=1}^{\infty} E(f_n) = \sum_{n=1}^{\infty} (1-\theta)^{n-1} = \frac{1}{\theta}$$

The variance of the number of times in state 0 is

(4) $$\mathrm{Var}(n_0) = \frac{1}{\theta}\,[\frac{2}{\theta} - 1] - \frac{1}{\theta^2} = \frac{1-\theta}{\theta^2}$$

For each item we define a sequence of response random variables, x_n, which take on the value 1 if an error occurs on trial n and which take on the value 0 if a success occurs on n. From the axioms, we have

(5) $$\Pr\{x_n = 1 | C_{1,n}\} = 0 \qquad \text{and} \qquad \Pr\{x_n = 1 | C_{0,n}\} = 1 - \frac{1}{N}$$

To obtain the probability of an error on trial n, q_n, we have

(6) $$q_n = \Pr\{x_n = 1\} = \Pr\{x_n = 1 | C_{1,n}\} \Pr\{C_{1,n}\} + \Pr\{x_n = 1 | C_{0,n}\} \Pr\{C_{0,n}\}$$
$$= 0 + (1 - \frac{1}{N})\ (1-\theta)^{n-1} = (1 - \frac{1}{N})\ (1-\theta)^{n-1}$$

The expected total number of errors, u_1, before perfect learning is given by

(7a) $$u_1 = E[\sum_{n=1}^{\infty} x_n] = \sum_{n=1}^{\infty} \Pr\{x_n = 1\} = \sum_{1}^{\infty} (1 - \frac{1}{N})\ (1-\theta)^{n-1} = \frac{1-\frac{1}{N}}{\theta}$$

An alternative way to derive this result is to find the average number of trials in state C_0, which is given by $1/\theta$ as in (3), and multiply by $1 - 1/N$ since this is the probability of an error for every trial that the process is in C_0.

In the expression for u_1, all errors are weighted equally. It is also possible to derive expressions for various weighted sums of errors, as Bush and Sternberg have shown for the linear model. The results here are exactly the same as their results. Three examples of the expectation of weighted error sums are given below:

(7b) $$E[\sum_{n=1}^{\infty} n x_n] = \sum_{n=1}^{\infty} n(1 - \frac{1}{N})(1-\theta)^{n-1} = \frac{1 - \frac{1}{N}}{\theta^2} = \frac{u_1}{\theta}$$

(7c) $$E\left[\sum_{n=1}^{\infty} \frac{x_n}{n}\right] = \sum_{n=1}^{\infty} \frac{(1 - \frac{1}{N})(1-\theta)^{n-1}}{n} = \frac{1 - \frac{1}{N}}{1-\theta} \log \frac{1}{\theta}$$

(7d) $$E\left[\sum_{n=1}^{\infty} \frac{x_n}{(n-1)!}\right] = (1 - \frac{1}{N}) \sum_{m=o}^{\infty} \frac{(1-\theta)^m}{m!} = (1 - \frac{1}{N})\, e^{1-\theta}$$

We next consider the variance of the total errors. However before proceeding to the variance problem we should state a theorem about the "autocorrelation", $c_{k,n}$, between errors on trial n and errors on trial $n+k$, $k = 1, 2, 3, \ldots$. We define $c_{k,n}$ as $x_n \cdot x_{n+k}$ and this will have a value of 1 only if errors occur on both trials n and n+k.

The expectation of $c_{k,n}$ is

$$(8) \qquad E(c_{k,n}) = E(x_n \cdot x_{n+k}) = E(x_{n+k}|x_n) \cdot E(x_n)$$

$$= \Pr\{x_{n+k} = 1|x_n = 1\} \Pr\{x_n = 1\} .$$

To find the conditional probability above, we merely note that for an error to occur on trial n+k it must be the case that conditioning has failed to occur during the intervening k trials, and moreover that the subject guesses incorrectly on trial n+k. The probability of this joint event is

$$(9) \qquad \Pr\{x_{n+k} = 1|x_n = 1\} = (1-\theta)^k \left(1 - \frac{1}{N}\right)$$

Therefore, we have

$$(10) \qquad E(c_{k,n}) = \left(1 - \frac{1}{N}\right) (1-\theta)^k \left(1 - \frac{1}{N}\right) (1-\theta)^{n-1}$$

A convenient statistic for comparison with data is obtained by taking the autocorrelation of x_n and x_{n+k} over all trials n of the experiment. We define c_k to be the mean value of this random variable and we have the following relation:

$$(11) \quad c_k = E\left[\sum_{n=1}^{\infty} x_n \, x_{n+k}\right] = \sum_{n=1}^{\infty} c_{k,n} = u_1\left(1 - \frac{1}{N}\right) (1-\theta)^k$$

for $k = 1, 2, 3, \ldots$.

We are now in a position to prove a theorem about the variance of the total errors.

Theorem:

(12) $$\mathrm{Var}[\sum_{n=1}^{\infty} x_n] = \frac{u_1}{2-\theta} [\frac{1}{N} + (1-\theta) (1 + 2u_1)]$$

where u_1 is the expected total errors as given in Equation 7a.

Proof:

(13) $$\mathrm{Var}[\sum_{n=1}^{\infty} x_n] = \sum_{n=1}^{\infty} \mathrm{Var}(x_n) + 2 \sum_{1\leq i<j} \mathrm{Cov}(x_i,x_j)$$

To get the first term, we note

(14) $$\mathrm{Var}(x_n) = q_n(1-q_n) = (1 - \frac{1}{N})(1-\theta)^{n-1} [1- (1- \frac{1}{N})(1-\theta)^{n-1}]$$

and assuming this over trials, n, we obtain

(15) $$\sum_{n=1}^{\infty} \mathrm{Var}(x_n) = \frac{1 - \frac{1}{N}}{\theta} - \frac{(1 - \frac{1}{N})^2}{1 - (1-\theta)^2} = u_1 \frac{(\frac{1}{N} + 1-\theta)}{2-\theta}$$

To get the second term, the covariance sum, we first note

(16) $$\mathrm{Cov}(x_i,x_j) = E(x_i,x_j) - E(x_i)\, E(x_j) .$$

The first term of (16) we have already calculated in (10), viz., for $j > i$

$$(17)\quad E(x_i, x_j) = (1 - \tfrac{1}{N})\,(1-\theta)^{j-i}\,(1 - \tfrac{1}{N})\,(1-\theta)^{i-1} = (1 - \tfrac{1}{N})^2\,(1-\theta)^{j-1}$$

So the covariance term is seen to be

$$(18)\quad \mathrm{Cov}(x_i, x_j) = (1 - \tfrac{1}{N})^2\,(1-\theta)^{j-1} - (1 - \tfrac{1}{N})\,(1-\theta)^{j-1}\,(1 - \tfrac{1}{N})(1-\theta)^{i-1}$$

$$= (1 - \tfrac{1}{N})^2\,(1-\theta)^{j-1}\,[1 - (1-\theta)^{i-1}]$$

The final task is to perform the summation over all trials i and j with $i < j$.

$$(19)\quad \sum_{i<j} \mathrm{Cov}(x_i, x_j) = \sum_{j=2}^{\infty} \sum_{i=1}^{j} (1 - \tfrac{1}{N})^2\,(1-\theta)^{j-1}\,[1 - (1-\theta)^{i-1}]$$

$$= (1 - \tfrac{1}{N})^2 \sum_{j=2}^{\infty} j(1-\theta)^{j-1} - (1 - \tfrac{1}{N})^2 \sum_{j=2}^{\infty} (1-\theta)^{j-1} \sum_{i=1}^{j} (1-\theta)^{i-1}$$

Performing these summations yields

$$(20)\quad \sum_{i<j} \mathrm{Cov}(x_i, x_j) = (1 - \tfrac{1}{N})^2\,\frac{(1-\theta)^2}{\theta^2} - (1 - \tfrac{1}{N})^2\,\frac{(1-\theta)}{\theta^2} + \frac{(1 - \frac{1}{N})^2\,(1-\theta)^3}{\theta[1-(1-\theta)^2]}$$

With algebraic manipulations, this can be simplified to

$$(21)\qquad \sum_{i<j} \mathrm{Cov}(x_i, x_j) = \frac{(1-\theta)}{2-\theta}\,u_1^2\ .$$

Putting together our results in (15) and (21), we obtain

$$(22)\quad \mathrm{Var}(\Sigma x_n) = \frac{u_1}{2-\theta}\left[1 - \theta + \frac{1}{N}\right] + \frac{2u_1^2}{2-\theta}(1-\theta) = \frac{u_1}{2-\theta}\left[\frac{1}{N} + (1-\theta)(1+2u_1)\right]$$

This concluses the proof.

III. Theorems about runs of errors.

Predictions about sequential features of the data may be obtained by considering runs of errors. To date only mean values of the various run distributions have been derived and higher moments will not be discussed. We let r_j represent the number of error runs of length j in an infinite number of trials and we seek the expectation of r_j. For these purposes, it is convenient to define another random variable, u_j, which counts the number of j-tuples of errors that occur in an infinite sequence of trials. Formally, we define u_j as

$$(23)\qquad u_j = \sum_{n=1}^{\infty} x_n\, x_{n+1} \cdots x_{n+j-1} \qquad \text{for} \quad j = 1, 2, \ldots$$

The product, $x_n\, x_{n+1} \cdots x_{n+j-1}$, has the value 1 only when j consecutive errors occur starting with the error on trial n. It may be seen that u_1 is just the total number of errors.

To make clear how the u_j are being counted and their relation to the r_j, consider the possible sequence

1111100110001101000 ... (all the rest zeros)

For this sequence, we would have

$u_1 = 10$ $u_2 = 6$ $u_3 = 3$ $u_4 = 2$ $u_5 = 1$

$r_1 = 1$ $r_2 = 2$ $r_3 = r_4 = 0$ $r_5 = 1$

and total number of error runs, $R = \sum_j r_j = 4$. In his excellent article on "Sequential Properties of Linear Models" [4], Bush has shown that the r_j can be expressed as linear combinations of the u_j; in particular,

(24) $$r_j = u_j - 2u_{j+1} + u_{j+2}$$

and

(25) $$R = \sum_{j=1}^{\infty} r_j = u_1 - u_2$$

The reader may verify these relations from the example given above; readers interested in the derivation of (24) and (25) should consult Bush's paper.

Having expressed the r_j in terms of the u_j, we now turn to deriving from the model the expected value of u_j. We proceed as follows:

(26) $$E(u_j) = E[\sum_{n=1}^{\infty} x_n \cdot x_{n+1} \cdots x_{n+j-1}]$$

$$= \sum_{n=1}^{\infty} \Pr\{x_n=1\} \Pr\{x_{n+1}=1|x_n=1\} \Pr\{x_{n+2}=1|x_n \cdot x_{n+1}=1\}$$

$$\cdots \Pr\{x_{n+j-1}=1|x_n \cdot x_{n+1} \cdots x_{n+j-2}=1\}$$

Because of the Markovian properties of our model, the lengthy conditional probabilities on the right-hand side can be simplified, viz.,

(27) $$\Pr\{x_{n+i}=1|x_n=1, x_{n+1} = 1, \ldots , x_{n+i-1} = 1\} = \Pr\{x_{n+i}=1|x_{n+i-1}= 1\}$$

That is, if the subject made an error on the preceding trial, then that's all the information there is to be extracted from the entire preceding sequence of responses. His error tells us that his conditioning state on the preceding trial was C_0 ; the probability of an error on the current trial is then

(28) $$\Pr\{x_{n+1} = 1|x_n = 1\} = \theta \cdot 0 + (1-\theta) (1 - \frac{1}{N}) = (1-\theta) (1 - \frac{1}{N}) = \alpha$$

and, moreover, this holds for any trial number n. Thus, using relations (27) and (28) we may rewrite our equation for u_j as follows:

(29) $$E(u_j) = \sum_{n=1}^{\infty} \Pr\{x_n=1\} \Pr\{x_{n+1}=1|x_n=1\} \cdots \Pr\{x_{n+j-1}=1|x_{n+j-2}=1\}$$

$$= \sum_{n=1}^{\infty} \Pr\{x_n= 1\} \underbrace{\alpha \cdot \cdots \cdot \alpha}_{(j-1) \text{ times}}$$

$$E(u_j) = \alpha^{j-1} \sum_{n=1}^{\infty} \Pr\{x_n=1\} = u_1 \alpha^{j-1}$$

With these values in hand we may now calculate R and r_j using relations (24) and (25). We obtain

(30) $$E(R) = E(u_1) - E(u_2) = u_1(1-\alpha)$$

(31) $$E(r_j) = E(u_j) - 2E(u_{j+1}) + E(u_{j+2}) = u_1(1-\alpha)^2 \alpha^{j-1}$$

$$= R(1-\alpha) \alpha^{j-1}$$

The average length of runs of errors is given by

(32) $$\text{ave. run length} = \frac{u_1}{R} = \frac{1}{1-\alpha}$$

It is also possible to derive trial-weighted k-tuples. One example is shown here where doublets of errors are weighted by the trial number of the leading error:

(33) $$E[\sum_{n=1}^{\infty} n\, x_n\, x_{n+1}] = \sum_{n=1}^{\infty} n \Pr\{x_n=1\} \Pr\{x_{n+1}=1|x_n=1\} = \frac{\alpha u_1}{\theta}$$

Another useful statistic in summarizing sequential properties of the data is obtained by taking the difference between the number of errors following errors and the number of errors following successes. Sternberg [5] has used this statistic in analyzing perseveration in response tendencies. Following Sternberg, we define B_n as

(34) $$B_n = x_{n+1}\, x_n - x_{n+1}(1-x_n)$$

B_n will have the value 1, 0 or -1. The expectation is

(35) $$E(B_n) = \Pr\{x_{n+1}=1|x_n=1\}\ \Pr\{x_n=1\} - \Pr\{x_{n+1}=1|x_n=0\}\ \Pr\{x_n=0\}$$

where

$$\Pr\{x_{n+1}=1|x_n=1\} = (1-\theta)\ (1 - \frac{1}{N}) = \alpha$$

and

$$\Pr\{x_{n+1}=1|x_n=0\} = \frac{\Pr\{x_{n+1}=1 \cap x_n=0\}}{\Pr\{x_n=0\}} = \frac{\frac{1}{N}\ (1-\theta)^{n-1}\ [(1-\theta)\ (1 - \frac{1}{N})]}{1 - (1 - \frac{1}{N})(1-\theta)^{n-1}}$$

so

$$E(B_n) = \alpha(1 - \frac{1}{N})(1-\theta)^{n-1} - \frac{\alpha}{N}\ (1-\theta)^{n-1} = \alpha(1-\theta)^{n-1}\ (1 - \frac{2}{N})$$

which, for $N > 2$ is always positive; that is, errors are more likely to follow errors than to follow successes. The sum of B_n over trials, B, has expectation

(36) $$E[B] = E[\sum_{n=1}^{\infty} B_n] = \frac{\alpha}{\theta}\ (1 - \frac{2}{N})$$

Similarly, we may define the number of alternations of successes and failures as

(37) $$A_n = x_{n+1}(1-x_n) + (1-x_{n+1})\, x_n$$

The expectation of A_n is

(38) $$E(A_n) = \frac{\alpha}{N}\,(1-\theta)^{n-1} + (1-\alpha)\,(1 - \frac{1}{N})\,(1-\theta)^{n-1}$$

And the average of the sum of A_n over trials is

(39) $$E[\sum_{n=1}^{\infty} A_n] = u_1[\theta + \frac{2(1-\theta)}{N}] \quad .$$

IV. Errors during various parts of learning.

In this section we derive the distribution of the number of errors before the first and before the second success and the distribution of the trial number on which the last error occurs. We define F as a random variable representing the number of errors before the first success. The possible values of F are $0, 1, 2, \ldots$. The probability that $F = 0$ is just the likelihood of guessing correctly on the first trial, which is $1/N$. For F to be $k > 0$, we require an error on the first trial, with probability $1 - 1/N$, $k-1$ successive failures without conditioning occurring, with probability $(1 - 1/N)^{k-1}\,(1-\theta)^{k-1}$, and following the k^{th} error, a success with probability $\theta + \frac{1}{N}\,(1-\theta)$.

Thus, the distribution of F is given by

$$(40) \qquad \Pr\{F = k\} = \begin{cases} \frac{1}{N} & \text{if} \quad k = 0 \\ (1 - \frac{1}{N})\ (1-\alpha)\ \alpha^{k-1} & \text{if} \quad k \geq 1 \end{cases}$$

where $\alpha = (1 - \frac{1}{N})\ (1-\theta)$ as before. The mean and variance of the number of errors before the first success are given by

$$(41) \qquad E(F) = \frac{1 - 1/N}{1 - \alpha}, \qquad Var(F) = E(F)\ [1 + (1-2\theta)\ E(F)]$$

Now define S as the number of errors before the second success. The probability that $S = 0$ is the likelihood of a correct guess on trial 1 and a correct response in trial 2, which is $\frac{1}{N}\ [\theta + \frac{(1-\theta)}{N}]$.

For general $k > 0$, the event of k errors before the second success can happen in three ways: (a) all errors through the first k trials with no conditioning, then the item becomes conditioned on the k^{th} trial and the second success occurs on trial k+2--the probability of this event being $(1 - \frac{1}{N})^k\ (1-\theta)^{k-1}\theta$; (b) the item becomes conditioned at the end of trial k+1 after having one success some place during the first k+1 trials--the probability of this event is $\frac{(k+1)}{N}\ (1 - \frac{1}{N})^k\ (1-\theta)^k\ \theta$; or (c) the second success occurs by guessing on trial k+2 and the first success occurred somewhere during the first k+1 trials, no conditioning occurring during these first k+1 trials--the probability of this event is $(\frac{1}{N})^2 (k+1)\ (1 - \frac{1}{N})^k\ (1-\theta)^{k+1}$. Putting together these considerations and simplifying, we arrive at the distribution of the number of errors before the second success:

(42) $$\Pr\{S = k\} = \begin{cases} s_o = \frac{1-\alpha}{N} & \text{for } k = 0 \\ (1 - \frac{1}{N})[\theta + s_o(1-\theta)(k+1)]\ \alpha^{k-1} & \text{for } k \geq 1 \end{cases}$$

The expected value of S is

(43) $$E(S) = \frac{1 - 1/N}{1 - \alpha} + \frac{\alpha}{N(1-\alpha)^2} = E(F) + \frac{\alpha}{N(1-\alpha)^2}$$

Thus, we see that the average errors before the second success can be expressed as the number of errors before the first success, $E(F)$, plus a term giving the expected number of errors intervening between the first and second success. It should be noted that the average trial number on which the first success occurs is $E(F) + 1$ and the trial number of the second success is $E(S) + 2$. The variance terms would be the same. The variance of S, the number of errors before the second success, may be obtained directly from the distribution in (42). After simplification, the result is

(44) $$\mathrm{Var}(S) = \frac{1}{1-\alpha} + \frac{1}{N(1-\alpha)^3}\,[4\alpha^2 + (1-\alpha)\ (2\alpha N-1)]$$

Let us now consider the total number of errors following the first success. If the first correct response came about through conditioning on the preceding trial, then we expect no further errors on that item; if, however, the first correct response came about by guessing, then we expect a few more errors. Let e represent the average number of subsequent errors following first successes which came about by guessing. The value of e may be derived to be

$$(45) \qquad e = \theta \cdot 0 + (1-\theta) \sum_{n=1}^{\infty} (1 - \tfrac{1}{N}) \ (1-\theta)^{n-1} = (1-\theta) \ u_1$$

If we let g represent the proportion of first correct responses which occur by guessing, then the mean number of errors following the first correct response (call it w) will be

$$(46) \qquad w = g \cdot e + (1-g) \cdot 0 = g \cdot e$$

That is, a proportion $1-g$ of the first correct responses occurred because of conditioning on the preceding trial and for these items we expect no subsequent errors. To complete our derivation we need to know g, the probability that the first correct response comes about by guessing. This result is easily derived as

$$(47) \qquad g = \frac{1}{N} + (1 - \frac{1}{N})(1-\theta) \ \frac{1}{N} + (1 - \frac{1}{N})^2 \ (1-\theta)^2 \ \frac{1}{N} + \cdots$$

$$= \frac{1}{N} \sum_{j=0}^{\infty} \alpha^j = \frac{1}{N(1-\alpha)}$$

Combining this result with (45) we obtain for mean errors following the first correct response,

$$(48) \qquad w = g \cdot e = \frac{(1-\theta) \ u_1}{N(1-\alpha)} = \frac{\alpha}{N\theta(1-\alpha)}$$

This result may be checked by subtracting the average errors to the first success from the average total errors, since we should get the same answer for w, viz.,

$$(49) \qquad w \equiv u_1 - E(F_1) = u_1 - \frac{(1 - \frac{1}{N})}{1 - \alpha} = \frac{\alpha}{N\theta(1-\alpha)}$$

The values of g and e derived above cannot be compared with the data because we can't tell whether a given first correct response has come about by conditioning or by guessing. One might think that 1-g could be estimated by the proportion of first correct responses which were followed by all correct on further trials. Although it is true that this "no more errors" property will hold for items whose first correct response occurred via prior conditioning, it is also true that some items whose first correct response occurred by guessing will also exhibit no further errors. In fact, it is possible to derive an expression for the proportion of sequences exhibiting this property of no errors following the first correct response. Let b represent the probability that an item will exhibit no further errors after its first correct response came about by guessing. We may derive the relation

$$(50) \qquad b = \theta + (1-\theta)\frac{1}{N}\theta + (1-\theta)^2\left(\frac{1}{N}\right)^2\theta + \cdots$$

$$= \theta \sum_{j=0}^{\infty} \left[\frac{(1-\theta)}{N}\right]^j = \frac{\theta}{1 - \frac{1-\theta}{N}}$$

That is, with probability θ the item was conditioned on the trial on which the first correct guess occurred; with probability $1-\theta$ conditioning did not occur on that trial, the S guessed correctly on the next trial with probability $1/N$ and the item became conditioned then with probability θ, and so on.

Using this result, we can now write an expression for the proportion, p, of items whose first correct response is followed by all correct responses. This includes a proportion $1-g$ whose first correct response came about because of prior conditioning, and a proportion $g \cdot b$ whose first correct response occurred by guessing but no further errors occurred.

(51) $$p = 1 - g + gb$$

$$= 1 - g(1-b) = 1 - \frac{(1-\theta)\ (N-1)}{[1 + (N-1)\theta]\ [\theta + N - 1]}$$

This proportion, p, of sequences characterized by no errors following the first success can be remarkably large. For example, if $\theta = .20$ and $N = 8$, we predict that 67.5% of the sequences will have no errors following the first correct response. The probability that no errors occur with a particular item is

(52) $$\Pr\{u_1 = 0\} = \frac{b}{N} = \frac{\theta}{\theta + N - 1}$$

That is, with probability $1/N$ the first guess is correct and with probability b no further errors are made. Thus, with $\theta = .20$, we expect the proportion of items displaying no errors to be about 9% when $N = 3$ and about 3% when $N = 8$.

Another interesting and exacting test of the Markov model arises when we ask the question: given that an error occurred on trial n, what is the expected number of errors that will follow before perfect learning? The rather surprising answer to this question according to the model is that the expected number of errors following an error is a constant independent of the trial number on which the leading error occurred. That is, if we observe an error on trial 1, the expected number of subsequent errors is predicted to be the same as that following an error that occurs on, say, trial 25. This result is counter to intuition but nonetheless follows strictly from the Markov model. The point of the matter is that if we observe an error on trial n, then we know the item was not conditioned prior to that trial; hence, we can assume our Markov process "starts" in state 0 on trial n and the expected number of subsequent errors is $(1-\theta)\ u_1$, as was derived for e in Equation 45. When an error occurs we may, so to speak, set the clock back to trial 1 as far as the model is concerned in predicting future behavior. Thus, the expected number of errors following an error on trial n is a constant independent of n. It should be noted that this prediction is markedly different from that delivered by the linear model which would predict that errors following an error on trial n is a monotone decreasing function of n.

As a preliminary test of these predictions, data from one group of 14 Ss learning a list of 10 paired associates were analyzed. For these purposes trials 2, 3, 4 and 5 were selected, and the average number of subsequent errors was determined for those item-subjects showing an error on trial 2, on trial 3, etc. The results, in terms of average subsequent errors, are 2.08 following an error occurring on the second trial, 2.03 following an error occurring on the third trial, and 2.05 following an error at the fourth or fifth trial (these were pooled to get a larger sample). Pooling the data for errors on trial 2 and trial 3, the mean is 2.05 which is identical with the average errors following errors that occur on the fourth or fifth trials. Thus, these limited results give rather striking confirmation to the "constancy" prediction from the one-element Markov model. However, more data analysis is required before we can make a confident decision in favor of the Markov model and reject the linear model.

The results obtained so far are useful in helping us to derive the distribution of the trial number on which the last error occurs in an infinite sequence of trials. For these purposes, let us define a random variable, n', which represents the trial number on which the last error occurs. Also we allow n' to be zero in case no errors occur at all. The probability distribution of n' is given by

$$\Pr\{n' = k\} = \begin{cases} \frac{b}{N} & \text{for} \quad k = 0 \\ b(1 - \frac{1}{N})\ (1-\theta)^{k-1} & \text{for} \quad k \geq 1 \end{cases} \tag{53}$$

where

$$b = \frac{\theta}{1 - \frac{(1-\theta)}{N}}$$

The first value for $k = 0$ has already been derived in (52). To obtain an error on trial k, it must be that conditioning has failed to occur on the preceding k-1 trials and the subject guesses incorrectly on trial k, and then makes no more errors with probability b. The mean and variance of n' are

$$E(n') = m = \frac{b(1 - \frac{1}{N})}{\theta^2} = \frac{b}{\theta} u_1 \tag{54}$$

$$\mathrm{Var}(n') = m[\ \frac{2}{\theta} - 1 - m]$$

We may use (54) to derive numerous relations between mean errors at different parts of learning. One example is the average number of trials between the first success and last failure; this is

$$E(n') - E(F + 1) = \frac{b}{\theta} u_1 - \frac{(1 - \frac{1}{N})}{1 - \alpha} - 1 \tag{55}$$

Similarly, the average number of trials between the second success and the last error is

$$(56) \qquad E(n') - E(S + 2) = \frac{bu_1}{\theta} - \frac{(1 - 1/N)}{1 - \alpha} - \frac{\alpha}{N(1-\alpha)^2} - 3$$

What we have accomplished in this section is a rather precise partitioning of the response sequence around three nodes--the first success, the second success, and the last failure. To illustrate this partitioning, Table 1 shows the various error quantities for the illustrative values of $N = 3$ and $\theta = .10$.

TABLE 1

Expected Error Quantities For the Values $N = 3$, $\theta = .10$.

1.	Total errors	6.67
2.	Errors before first success	1.67
3.	Errors between first and second success	1.25
4.	Errors before second success	2.92
5.	Error following 1^{st} success	5.00
6.	Errors following 2^{nd} success	3.75
7.	Trial of last error	9.53
8.	Trials from first success to last error	6.86
9.	Trials from second success to last error	4.61
10.	Proportion of items with zero errors	4.76%
11.	Proportion of items with no errors following first success	28.58%

V. Statistics across items within a particular trial.

The preceding analysis has been carried out for single items across trials. In applying this model to paired-associate learning, experimental conditions can be arranged so that we may assume that each of the K items in the list has the same learning rate constant, θ. That is, the stimulus items may be so chosen that all K items may be considered as homogeneous in difficulty. To the extent that these homogeneous conditions are satisfied, we may consider a run through the list (a trial) as providing a sample of size K from a binomial population. To make this notion explicit, let us arbitrarily denote the response to the first item in the list on trial n by a random variable $x_{1,n}$, the response to the second item of the list by $x_{2,n}$, ... , the response to the K^{th} and last item by $x_{k,n}$. Because we usually randomize the order of the items between trials, the first item on trial n+1 will not usually be the same as the first item on trial n. Nonetheless, $x_{1,n+1}$ would still be defined as the response random variable for the first item on trial n+1, whatever that item may be. The reasoning here is that if we consider the response to each item as being one sample from the same binomial population, then it makes no difference in which order the samples are drawn or labelled.

The response random variable for the i^{th} item in the list is defined as

$$(57)\qquad x_{i,n} = \begin{cases} 1 & \text{if error on } i^{th} \text{ item on trial } n \\ 0 & \text{if correct on } i^{th} \text{ item on trial } n \end{cases}$$

and the parameter of the distribution is

$$(58)\qquad \Pr\{x_{i,n} = 1\} = q_n = (1 - \tfrac{1}{N})(1-\theta)^{n-1}$$

We define the total number of errors on trial n, T_n, as the sum of $x_{i,n}$ over the list of K items:

$$(59)\qquad T_n = \sum_{i=1}^{K} x_{i,n}$$

It is clear that T_n has the binomial distribution given by

$$(60)\qquad \Pr\{T_n = j\} = \binom{K}{j} q_n^j (1-q_n)^{K-j}$$

with mean and variance

$$(61)\qquad E(T_n) = Kq_n, \qquad Var(T_n) = K\, q_n(1-q_n)$$

Using (60) and Equation 58 for q_n we can calculate the probability that, say, exactly half the items are correct on trial n (for K even, of course). This probability is

$$(62)\quad \Pr\{T_n = \frac{K}{2}\} = \binom{K}{K/2} [(1 - \frac{1}{N})(1-\theta)^{n-1}]^{K/2} [1 - (1 - \frac{1}{N})(1-\theta)^{n-1}]^{K/2}$$

$$= \binom{K}{K/2} [(1 - \frac{1}{N})(1-\theta)^{n-1} - (1 - \frac{1}{N})^2 (1-\theta)^{2(n-1)}]^{K/2}$$

$$= \binom{K}{K/2} \sigma_{P_n}^K$$

If N is large, say 8, then this probability begins at a low value, increases to a maximum with trials, then drops eventually to zero as n becomes large. As n becomes large, all items become learned and the probability of getting exactly half the items correctly decreases to zero.

One can also use (60) and the Equation 58 for q_n to calculate other statistics of interest such as the probability of getting at least j correct out of the K items. Such probabilites are obviously related to "criterion of learning" scores. Consider the criterion of all items correct: What is the probability of getting all K items correct on trial n? This probability is

$$(63)\quad \Pr\{T_n = K\} = [1 - (1 - \frac{1}{N})(1-\theta)^{n-1}]^K = p_n^K$$

This probability can be used to compute the distribution of the number of trials before a criterion of K correct is met. Let C be the number of trials before meeting the criterion of one correct run through the list. Then the distribution of C is

$$
\begin{aligned}
&\Pr\{C = 1\} = p_1^K \\
&\Pr\{C = 2\} = (1 - p_1^K)\, p_2^K \\
(64) \qquad &\vdots \\
&\Pr\{C = n\} = p_n^K \prod_{j=1}^{n-1} (1 - p_j^K)
\end{aligned}
$$

where

$$p_j = 1-(1 - \tfrac{1}{N})\ (1 - \theta)^{j-1}$$

The mean trials to criterion would then be the expectation of the random variable defined by (64). Unfortunately, no analtyic solution to the problem is available because of the product term on the right side of (64). However, for reasonable values of θ, $E(C)$ could be approximated numerically from the first ten or so terms of the first moment sum.

A final issue to be considered concerns run statistics within a trial. Analogous to our definition of u_j in (23) for a single item sequence, we may define

$$(65) \qquad y_{j,n} = \sum_{i=1}^{K-(j-1)} x_{i,n}\, x_{i+1,n} \cdots x_{i+j-1,n}$$

The $y_{j,n}$ count j-tuples of errors on the n^{th} trial. Because we have assumed that the $x_{i,n}$ and $x_{j,n}$ are independent samples from the binomial population, we may write the expectation of $y_{j,n}$ as

$$(66) \qquad E(y_{j,n}) = \sum_{i=1}^{K-(j-1)} \underbrace{q_n \cdot q_n \cdots q_n}_{j \text{ times}} = (K + 1 - j)\, q_n^j$$

Using the $y_{j,n}$ we may define runs of j-errors within trial n--call it $m_{j,n}$-- analogously to (24);

$$(67) \qquad m_{j,n} = y_{j,n} - 2y_{j+1,n} + y_{j+2,n}$$

The expectation of $m_{j,n}$ is then

$$(68) \qquad E(m_{j,n}) = q_n^j(1 - q_n)^2\ (K + 1 - j) + 2q_n^{j+1}(1 - q_n)$$

Since q_n is a decreasing function of n, the limit of $E(m_{j,n})$ as n increases is zero. The total number of runs of errors on trial n is

$$(69) \qquad \hat{R}_n = y_{1,n} - y_{2,n} = K\, q_n - (K-1)\, q_n^2$$

The $m_{j,n}$ and $\hat{R}_n$ so calculated are for a single trial. To get a larger sample we may sum these terms over trials of an infinite sequence. There is no problem here with convergence of these sums because eventually all items are answered correctly and runs of errors would cease occurring. We may thus define m_j as the sum over all trials of $m_{j,n}$. The expectation of m_j is seen to be

$$(70) \quad E(m_j) = E[\sum_{n=1}^{\infty} m_{j,n}] = \sum_{n=1}^{\infty} E(m_{j,n})$$

$$= \frac{(K+1-j)[1-\frac{1}{N}]^j}{1-(1-\theta)^j} - \frac{2(K-j)[1-\frac{1}{N}]^{j+1}}{1-(1-\theta)^{j+1}} + \frac{(K-1-j)[1-\frac{1}{N}]^{j+2}}{1-(1-\theta)^{j+2}}$$

Also, we may derive the total within-trial runs of errors to be

$$(71) \quad E(\hat{R}) = E[\sum_{n=1}^{\infty} \hat{R}_n] = \sum_{n=1}^{\infty} E(\hat{R}_n) = \frac{u_1}{2-\theta}[1 + K(1-\theta) + \frac{K-1}{N}]$$

Discussion.

The previous sections have illustrated some of the large number of theorems derivable from the one-element model. These theorems are stated in a form equating some property of the data (e.g., number of error runs) to a derived expression which, once N has been fixed by the experimenter, is only a function θ. One of these equations would be used to estimate θ and the remainder may be used to obtain numerical predictions about various measures of the data. The statistic to be used in estimating θ is arbitrary; the writer has usually used average errors per item since this statistic uses most of the data and is probably a minimum variance estimator.

As mentioned earlier, the one-element model developed here has proved quite adequate in accounting for the data from several experiments on paired-associate learning. It is indeed a pleasant coincidence in mathematical learning theory when we find a model that comes close to the data and for which many important predictions are derivable in closed form. With the mathematical analysis of the one-element model now in reasonably complete form, the machinery is available for making detailed comparisons of the linear model and the one-element model in their predictions of experimental results.

REFERENCES

[1] Bush, R. R. and Sternberg, S. H. A single-operator model. In Bush, R. R. and Estes, W. K. (Ed.): Studies in mathematical learning theory. Stanford University Press, 1959.

[2] Suppes, P. and Atkinson, R.C. Markov learning models for multiperson interactions, Stanford University Press, 1960.

[3] Estes, W. K. Components and pattern models with Markovian interpretations. In Bush, R. R. and Estes, W. K. (Ed.): Studies in mathematical learning theory. Stanford University Press, 1959.

[4] 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.

[5] Sternberg, S. H. Applications of four models to sequential dependence in human learning. In Bush, R. R. and Estes, W.K. (Ed.): Studies in mathematical learning theory. Stanford University Press, 1959.

STANFORD UNIVERSITY
TECHNICAL REPORT DISTRIBUTION LIST
CONTRACT Nonr 225(17)
(NR 171-034)

Armed Services Technical
Information Agency
Arlington Hall Station
Arlington 12, Virginia 10

Commanding Officer
Office of Naval Research
Branch Office
Navy No. 100, Fleet Post Office
New York, New York 35

Director, Naval Research Laboratory
Attn: Technical Information Officer
Washington 25, D. C. 6

Director, USAF Project RAND
Via: AF Liaison Office
The RAND Corp. Library
1700 Main St.
Santa Monica, Calif. 1

Office of Naval Research
Group Psychology Branch
Code 452
Department of the Navy
Washington 25, D. C. 5

Office of Naval Research
Branch Office
346 Broadway
New York 13, N. Y. 1

Office of Naval Research
Branch Office
1000 Geary St.
San Francisco 9, Calif. 1

Office of Naval Research
Branch Office
1030 Green St.
Pasadena 1, Calif. 1

Office of Naval Research
Branch Office
The John Crerar Library Bldg.
86 E. Randolph St.
Chicago 1, Ill. 1

Office of Naval Research
Logistics Branch, Code 436
Department of the Navy
Washington 25, D. C. 1

Office of Naval Research
Mathematics Division, Code 430
Department of the Navy
Washington 25, D. C. 1

Operations Research Office
6935 Arlington Road
Bethesda 14, Md.
Attn: The Library 1

Office of Technical Services
Department of Commerce
Washington 25, D. C. 1

The Logistics Research Project
The George Washington University
707 - 22nd St., N. W.
Washington 7, D. C. 1

The RAND Corporation
1700 Main St.
Santa Monica, Calif.
Attn: Dr. John Kennedy 1

Library
Cowles Foundation for Research
in Economics
Yale Station, Box 2125
New Haven, Conn. 1

Center for Philosophy of Science
University of Minnesota
Minneapolis 14, Minn. 1

Stanford Research Institute
Document Center
333 Ravenswood Ave.
Menlo Park, Calif. 1

Wright Air Development Center
Attn: WCLDPP, Mrs. Southern
Wright-Patterson AF Base, Ohio 1

Department of Mathematics
Michigan State University
East Lansing, Michigan 1

Professor Ernest Adams
Department of Philosophy
University of California
Berkeley 4, California 1

Professor Alan Ross Anderson
Department of Philosophy
Yale University
New Haven, Conn. 1

Professor Norman H. Anderson
Department of Psychology
University of California
Los Angeles 24, Calif. 1

Professor T. W. Anderson
Department of Statistics
Columbia University
New York 27, N. Y. 1

Dr. R. F. Bales
Department of Social Relations
Harvard University
Cambridge, Mass. 1

Professor Edward G. Begle
School Mathematics Study Group
Yale University
Drawer 2502A
New Haven, Conn. 1

Professor Gustav Bergmann
Department of Philosophy
State University of Iowa
Iowa City, Iowa 1

Professor Max Black
Department of Philosophy
Cornell University
Ithaca, New York 1

Professor David Blackwell
Department of Statistics
University of California
Berkeley 4, Calif. 1

Mr. Richard S. Bogartz
Psychology Department
U.C.L.A.
Los Angeles 24, Calif. 1

Professor Lyle E. Bourne, Jr.
Department of Psychology
University of Utah
Salt Lake City, Utah 1

Professor C. J. Burke
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor R. R. Bush
106 College Hall
University of Pennsylvania
Philadelphia 4, Pa. 1

Dr. Donald Campbell
Department of Psychology
Northwestern University
Evanston, Ill. 1

Mr. J. Merrill Carlsmith
Department of Social Relations
Harvard University
Cambridge 38, Mass. 1

Professor Rudolf Carnap
Department of Philosophy
University of California
Los Angeles 24, Calif. 1

Professor Edward C. Carterette
Department of Psychology
University of California
Los Angeles 24, Calif. 1

Professor Noam Chomsky
Department of Philosophy
Massachusetts Institute of
Technology
Cambridge, Massachusetts 1

Professor C. W. Churchman
School of Business Administration
University of California
Berkeley 4, California 1

Professor James Coleman
Department of Social Relations
Johns Hopkins University
Baltimore, Maryland 1

Dr. Clyde H. Coombs
Department of Psychology
University of Michigan
Ann Arbor, Michigan

Professor Gerard Debreu
Cowles Foundation for Research
in Economics
Yale Station, Box 2125
New Haven, Connecticut 1

Dr. J. A. Deutsch
Center for Advanced Study in
the Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, California 1

Professor Robert Dorfman
Department of Economics
Harvard University
Cambridge 38, Massachusetts 1

Professor Burton Dreben
Department of Philosophy
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor P. H. Dubois
Department of Psychology
Washington University
St. Louis 5, Missouri 1

Dr. Ward Edwards
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor W. K. Estes
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor Robert Fagot
Department of Psychology
University of Oregon
Eugene, Oregon 1

Professor Merrill M. Flood
231 West Engineering
University of Michigan
Ann Arbor, Michigan 1

Professor Raymond Frankmann
Department of Psychology
University of Illinois
Urbana, Illinois 1

Professor Milton Friedman
Department of Economics
University of Chicago
Chicago 37, Illinois 1

Dr. Eugene Galanter
Department of Psychology
University of Pennsylvania
Philadelphia 4, Pennsylvania 1

Professor Johan Galtung
Department of Sociology
Columbia University
New York 27, New York 1

Professor Wendell R. Garner
Department of Psychology
Johns Hopkins University
Baltimore 18, Maryland 1

Dr. Murray Gerstenhaber
Institute for Advanced Study
Department of Mathematics
Princeton, New Jersey 1

Professor Leo A. Goodman
Statistical Research Center
University of Chicago
Chicago 37, Illinois 1

Professor Nelson Goodman
Department of Philosophy
University of Pennsylvania
Philadelphia, Pennsylvania 1

Professor Harold Gulliksen
Educational Testing Service
20 Nassau St.
Princeton, New Jersey 1

Dr. Luther Haibt
International Business Machines Corporation
Yorktown Heights, New York 1

Professor Harold W. Hake
Department of Psychology
University of Illinois
Urbana, Illinois 1

Dr. Raphael Hanson
Department of Psychology
Reed College
Portland 2, Oregon 1

Professor Albert H. Hastorf
Department of Psychology
Dartmouth College
Hanover, New Hampshire 1

Professor Carl G. Hempel
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Davis H. Howes
Department of Economic and Social Science
Massachusetts Institute of Technology
Cambridge 39, Massachusetts 1

Professor Leonid Hurwicz
School of Business
University of Minnesota
Minneapolis 14, Minnesota 1

Professor Rheem F. Jarrett
Department of Psychology
University of California
Berkeley 4, California

Professor Lyle V. Jones
Department of Psychology
University of North Carolina
Chapel Hill, North Carolina 1

Professor Donald Kalish
Department of Philosophy
University of California
Los Angeles 24, California 1

Professor Harold H. Kelley
Laboratory for Research in Social Relations
University of Minnesota
Minneapolis 14, Minnesota 1

Professor John G. Kemeny
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Dr. William Kessen
Center for Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, California 1

Professor T. C. Koopmans
Cowles Foundation for Research in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Mr. Franklin Krasne
Department of Psychology
333 Cedar St.
Yale University
New Haven, Connecticut 1

Professor W. Kruskal
Department of Statistics
Eckart Hall 127
University of Chicago
Chicago 37, Illinois 1

Professor David La Berge
Department of Psychology
University of Minnesota
Minneapolis 14, Minnesota 1

Professor Paul Lazarfeld
Department of Sociology
Columbia University
New York 27, New York 1

Dr. Bernhardt Lieberman
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Frank Logan
Department of Psychology
Yale University
New Haven, Connecticut 1

Professor Duncan Luce
Department of Psychology
University of Pennsylvania
Philadelphia 4, Pennsylvania 1

Dr. David McConnell
RF 887
1314 Kinnear Rd.
Columbus 8, Ohio 1

Professor William J. McGill
Department of Psychology
Columbia University
New York 27, New York 1

Professor Robert McGinnis
Department of Sociology
University of Wisconsin
Madison 6, Wisconsin 1

Dr. Benoit Mandelbrot
Hawkes Ave.
Ossining, New York 1

Professor Jacob Marschak
Box 2125, Yale Station
New Haven, Connecticut 1

Professor Samuel Messick
Educational Testing Service
Princeton University
Princeton, New Jersey 1

Professor G. A. Miller
Department of Psychology
Harvard University
Cambridge 38, Massachusetts 1

Professor Richard C. Montague
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. O. K. Moore
Department of Sociology
Box 1965, Yale Station
New Haven, Connecticut 1

Professor Sidney Morgenbesser
Department of Philosophy
Columbia University
New York 27, New York 1

Professor Oscar Morgenstern
Department of Economics and Social Institutions
Princeton University
Princeton, New Jersey 1

Professor Frederick Mosteller
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Ernest Nagel
Center for Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, California 1

Professor Theodore M. Newcomb
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor Jerzy Neyman
Department of Statistics
University of California
Berkeley 4, California 1

Professor Vincent Nowlis
Department of Psychology
Rochester University
Rochester, New York 1

Professor A. G. Papandreou
Department of Economics
University of California
Berkeley 4, California 1

Dr. Juliet Popper
Department of Psychology
University of Kansas
Lawrence, Kansas 1

Professor William F. Prokasy, Jr.
Department of Psychology
Pennsylvania State University
University Park, Pennsylvania 1

Dean G. Pruitt
Political Science Department
Northwestern University
Evanston, Illinois 1

Professor Hilary Putnam
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Willard V. Quine
Department of Philosophy
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor Roy Radner
Department of Economics
University of California
Berkeley 4, California 1

Professor Howard Raiffa
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Dr. Anatol Rapoport
Mental Health Research Institute
University of Michigan
Ann Arbor, Michigan 1

Professor Nicholos Rashevsky
University of Chicago
Chicago 37, Illinois 1

Professor Philburn Ratoosh
Department of Psychology
Ohio State University
Columbus, Ohio 1

Professor Frank Restle
Department of Psychology
Michigan State University
East Lansing, Michigan 1

Dr. Henry Riecken, Director
Social Science Division
National Science Foundation
Washington 25, D. C. 1

Professor David Rosenblatt
American University
Washington 6, D. C. 1

Dr. Robert E. Ross
Electric Boat Division
General Dynamics Corporation
Groton, Connecticut 1

Alan J. Rowe, Manager
Business Management
Control Systems Research
System Development Corporation
2500 Colorado Ave.
Santa Monica, California 1

Professor Herman Rubin
Department of Statistics
Michigan State University
East Lansing, Michigan 1

Professor Richard S. Rudner
Department of Philosophy
Michigan State University
East Lansing, Michigan 1

Professor Paul Samuelson
Department of Economics
Massachusetts Institute of Technology
Cambridge, Massachusetts 1

Dr. I. Richard Savage
School of Business
Vincent Hall
University of Minnesota
Minneapolis, Minnesota 1

Professor L. J. Savage
Department of Statistics
University of Chicago
Chicago 37, Illinois

Dr. Dana Scott
Eckart Hall
University of Chicago
Chicago 37, Ill. 1

Dr. C. P. Seitz
Special Devices Center
Office of Naval Research
Sands Point
Port Washington
Long Island, N. Y. 1

Professor Marvin E. Shaw
Department of Psychology
University of Florida
Gainesville, Florida 1

Mr. R. L. Shuey, Manager
Research Laboratory
General Electric Company
P. O. Box 1088
Schenectady, N. Y. 1

Professor Sidney Siegel
Department of Psychology
Pennsylvania State University
University Park, Pa. 1

Professor Herbert Simon
Carnegie Institute of Technology
Schenley Park
Pittsburgh, Pa. 1

Professor J. L. Snell
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Professor K. W. Spence
Department of Psychology
State University of Iowa
Iowa City, Iowa 1

Professor S. Smith Stevens
Memorial Hall
Harvard University
Cambridge 38, Mass. 1

Dr. Donald W. Stilson
Department of Psychology
University of Colorado
Boulder, Colorado 1

Professor Marshall Stone
Department of Mathematics
University of Chicago
Chicago 37, Ill. 1

Dr. Dewey B. Stuit
108 Schaeffer Hall
State University of Iowa
Iowa City, Iowa 1

Professor John Swets
Psychology Section
Dept. of Econ. and Socl. Sci.
Mass. Institute of Technology
Cambridge 39, Mass. 1

Professor Alfred Tarski
Department of Mathematics
University of California
Berkeley 4, California 1

Professor G. L. Thompson
Department of Mathematics
Ohio Wesleyan
Delaware, Ohio 1

Dr. Robert L. Thorndike
Teachers College
Columbia University
New York 27, N. Y. 1

Professor R. M. Thrall
University of Michigan
Engineering Research Institute
Ann Arbor, Michigan 1

Dr. E. Paul Torrance
Bureau of Educational Research
University of Minnesota
Minneapolis 14, Minnesota 1

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, N. J. 1

Dr. Ledyard R. Tucker
Educational Testing Service
20 Nassau St.
Princeton, N. J. 1

Professor John W. Tukey
Fine Hall
Princeton, New Jersey 1

Professor John van Laer
Department of Psychology
Northwestern University
Chicago, Ill. 1

Professor Edward L. Walker
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. John T. Wilson
National Science Foundation
1520 H Street, N. W.
Washington 25, D. C. 1

Professor Kellogg Wilson
Kent State University
Kent, Ohio 1

Professor J. Wolfowitz
Department of Mathematics
Cornell University
Ithaca, N. Y. 1

Professor Robert J. Wolfson
Department of Economics
Michigan State University
East Lansing, Michigan 1

Professor David Zeaman
Department of Psychology
The University of Connecticut
Storrs, Conn. 1

Lt. Donald L. Zink
Engineering Psychology Branch
WCLDPPP
Aerospace Medical Lab.
Hq. Wright Air Dev. Center
Wright-Patterson AF Base, Ohio 1

<u>Distribution via ONR London</u>

Commanding Officer
Branch Office
Navy No. 100
Fleet Post Office
New York, New York

Professor Maurice Allais
15 Rue des Gates-Ceps
Saint-Cloud, (S.-O.)
FRANCE 1

Dr. R. J. Audley
Department of Psychology
University College
London WC 1, ENGLAND 1

Professor E. W. Beth
Bern, Zweerskade 23, I
Amsterdam Z.,
NETHERLANDS 1

Professor R. B. Braithwaite
Kings College
Cambridge, ENGLAND 1

Dr. F. Bresson
Laboratoire de Psychologie
Experimentale et Comparee
Ecole Pratique des Hautes Etudes
46 Rue St. Jacques
Paris (Ve.). FRANCE 1

Dr. John Brown
Department of Psychology
Birkbeck College
Malet Street
London WC 1, ENGLAND 1

Dr. Violet Cane
Newham College
Cambridge, ENGLAND 1

Dr. H.C.J. Duijker
Psychologisch Laboratorium
Keiszergracht
Amsterdam 613/C, NETHERLANDS 1

Mr. Michael Dummett
All Souls' College
Oxford, ENGLAND 1

Professor Jacques H. Dreze
5 Avenue Princesse Lydia
Heverle-Louvain, BELGIUM 1

Dr. Jean Engler
Department of Psychology
University of London
Gower St.
London, ENGLAND 1

Professor J. M. Faverge
Universite de Bruxelles
67, rue Franz Merjay
Brussels, BELGIUM 1

Dr. C. Flament
Laboratoire de Psychologie
Experimentale et Comparee
Ecole Pratique des Hautes Etudes
46 Rue St. Jacques
Paris (Ve), FRANCE 1

Professor Maurice Frechet
Institut H. Poincare
11 Rue P. Curie
Paris 5, FRANCE 1

Dr. I. J. Good
25 Scott House
Cheltenham, ENGLAND 1

Dr. T. T. ten Have
Sociaal - Paed. Instituut
Singel 453
Amsterdam, NETHERLANDS 1

Dr. W. E. Hick
Psychology Laboratory
Downing Place
Cambridge, ENGLAND 1

Institut fur Math. Logik
Universitat
Schlossplatz 2
Munster in Westfalen
GERMANY 1

Dr. A. R. Jonckheere
Department of Psychology
University College
London WC 1, ENGLAND 1

Mr. David Kendall
Magdalen College
Oxford, ENGLAND 1

Mr. E. J. Lemmon
Trinity College
Oxford, ENGLAND 1

Professor P. Lorenzen
Philosophisches Seminar
Der Universtät
Kiel, WEST GERMANY 1

Mr. Henri Rouanet
Centre d'Etudes et Recherches
Psychotechniques
13, rue Paul-Chautard
Paris (XVe), FRANCE 1

Madame Madeleine Schlag-Rey
Institut de Sociologie Solvay
Parc Leopold
Brussels, BELGIUM 1

Dr. Saul Sternberg
St. John's College
Cambridge, ENGLAND 1

Dr. N. S. Sutherland
Institute of Experimental Psychology
1 South Parks Road
Oxford, ENGLAND 1

Professor D. van Dantzig
Mathematical Centre
Statistical Department
2de Boerhaavestraat 49
Amsterdam, THE NETHERLANDS 1

Professor O. L. Zangwill
Psychological Laboratory
Downing Place
Cambridge, ENGLAND 1

Other Foreign Addresses

Professor Y. Bar-Hillel
Department for History and Philosophy
of Science
Hebrew University
Jerusalem, ISRAEL 1

L. Guttman
Israel Institute of Applied Social
Research
David Hamlech No. 1
Jerusalem, ISRAEL 1

Dr. John Harsanyi
Department of Social Philosophy
Australian National University
GPO Box 4, Canberra, A.C.T.
AUSTRALIA 1

Dr. Georg Karlsson
Sociologiska Institutionen
Uppsala, SWEDEN 1

Dr. T. Markkanen
Toironkatu 1.B36
Helsinki, FINLAND 1

Professor Hukukane Nikaido
The Institute of Social and
Economic Research
Osaka University
Toyonaka, Osaka
JAPAN 1

Dr. J. Pfanzagl
Institut Fur Statistik
Universitat Wien
Wien, AUSTRIA 1

Dr. Masanao Toda
Department of Experimental Psychology
Faculty of Letters
Hokkaido University
Sapporo, Hokkaido
JAPAN 1

Additional copies for project
leaders and assistants, office
file, and reserve for future
requirements 25