JOURNAL OF MATHEMATICAL PSYCHOLOGY: 6, 62–80 (1969)

A Three State Markov Model for Discrimination Learning[1]

DOMINIC W. MASSARO

University of California, San Diego, La Jolla, California 92037

Previous treatments of discrimination learning have been ineffective in predicting both marginal responding and reinforcement effects in two-stimulus probability learning. A three state Markov model is proposed which is designed to account for these results. The basic assumption of the theory is that the subject is reinforced for behaving appropriately to a distinctive cue rather than for making a specific motor response. Empirical results that include variations of π, β, and cue similarity are accounted for by the model.

In the present paper an attempt is made to quantify the view that discrimination learning is a form of nonspatial selective learning in which the subject learns to behave in relation to some particular set of stimuli. The basic assumption of this theory is that the subject is reinforced for behaving *appropriately* to a distinctive cue rather than for making a specific motor response. For simplicity, the discussion will be limited to two-stimulus, two-response problems, but the formulation can be extended readily to multi-stimulus response situations.

The present paper deals with a specific experimental paradigm. An experiment consists of a series of trials each of which commences with the onset of one of two stimuli, T_1 or T_2. The probability of T_1 is β, and the probability of T_2 is $1 - \beta$. Two responses, A_1 and A_2, predicting events E_1 and E_2, respectively, are available to the subject. On a T_1 trial, an A_1 response is correct with probability π_1, and an A_2 response is correct with probability $1 - \pi_1$. On a T_2 trial, an A_1 response is correct with probability π_2, and an A_2 response is correct with probability $1 - \pi_2$.

Previous treatments of discrimination learning within statistical learning theory formulations have assumed common stimulus elements between trial types (Atkinson, 1958; Burke and Estes, 1957). Both these common element models and a conditioning parameter model by Lee (1966) predict asymptotes of $P(A_1 \mid T_1)$ and $P(A_2 \mid T_2)$ that cannot exceed π_1 and $1 - \pi_2$, respectively. However, overshooting has been found

[1] This research was carried out during the author's tenure as a National Aeronautics and Space Administration predoctoral fellow at the University of Massachusetts under grant NSG(T)-137. The author would like to thank John W. Moore and Jerry L. Myers for their helpful comments during all phases of the present research.

in recent experiments in different laboratories (Massaro, Halpern, and Moore, 1968; Myers and Cruse, 1968). For example, when $\pi_1 = .85$, $P(A_1 \mid T_1)$ exceeded matching if $\pi_2 = .85$ or $.15$. However, when $\pi_2 = .5$, $P(A_1 \mid T_1) < \pi_1$. These results confirmed the findings of Popper and Atkinson (1958) that, for a fixed value of π_1, $P(A_1 \mid T_1)$ first decreased and then increased as π_2 went from π_1 to $1 - \pi_1$. These findings oppose the predictions of the Burke and Estes (1957) and Lee (1966) models which hold that the response probability on one trial type is a linear function of the event probabilities on the other trial type. The results seem to indicate that a discriminative stimulus with little or no cue value (i.e., one reinforced on a 50–50 basis) has a depression effect on $P(A_1 \mid T_1)$ compared to a stimulus with cue value (i.e., one reinforced on a 85–15 basis).

The above models also predict that the reinforcement of a response tends to make that *particular* response more likely on succeeding trials, not only for trials when the same stimulus is presented, but for both trial types. Therefore, both the Burke–Estes and Atkinson models predict

$$P(A_{1,n+1} \mid A_{1,n}E_{1,n}) \geqslant P(A_{1,n+1} \mid A_{i,n}E_{j,n}) \geqslant P(A_{1,n+1} \mid A_{2,n}E_{2,n}), \qquad i \neq j.$$

Massaro *et al.* (1968) and Myers and Cruse (1968) have also found first order conditional probabilitities that do not hold to the above rank-ordering. More specifically, under symmetrical reinforcement schedules (e.g., $\pi_1 = .8$, $\pi_2 = .2$), the rank-ordering found when the stimulus presented was the same as that presented on the previous trial ($T_{i,n+1} = T_{i,n}$) did agree with the predicted rank-ordering. However, the rank-ordering found when $T_{i,n+1} \neq T_{i,n}$ was exactly opposite that found when $T_{i,n+1} = T_{i,n}$. This finding seems to indicate that subjects are reinforced, not for making a specific motor response, but for making an *appropriate* response in the sense of having the highest likelihood of being correct on that particular stimulus trial. For example, if the right-hand light is the most frequent event given the trial type and the subject is reinforced for making a right-hand prediction, he is also reinforced for making an appropriate response and therefore he will be more likely to make the same response on the next trial provided it is of the same type. But if the next trial is different, the subject will still be more likely to make the appropriate response which means a response that is physically opposite from the response on the previous trial. Therefore, the Markov model presented in this paper views the subject as being reinforced not for a particular (right or left) event prediction, but for an appropriate (most frequently correct) or inappropriate response.

Variations of cue similarity in discrimination learning when $\pi_1 > .5$ and $\pi_2 < .5$ have been investigated only recently (Massaro *et al.*, 1968, Moore and Halpern, 1965). The Burke–Estes model (1957) predicts decreased dependence of $P(A_1 \mid T_1)$ on π_2 with decreases in cue similarity so that $P(A_1 \mid T_1)$ equals π_1 with sufficient cue distinctiveness. On the other hand, Atkinson's (1958) observing response model predicts no difference in asymptotic responding as a function of cue similarity when $\pi_1 = 1$

and $\pi_2 = 0$. Moore and Halpern did find that asymptotic responding was related to cue similarity, when $\pi_1 = 1, \pi_2 = 0$, such that $P(A_1 \mid T_1)$ approached $P(E_1)$ with increases in cue similarity. This result agrees with the Burke–Estes model but not the Atkinson model. However, Massaro *et al.*, employing values of $\pi_1 = .8$ and $\pi_2 = .8$, .5, or .2, found that $P(A_1 \mid T_1)$ did not become independent of π_2 with decreases in cue similarity. The $P(A_1 \mid T_1)$ was a linear function of π_2 when the trial types were highly confusable, independent of π_2 and equal to π_1 at an intermediate level of cue similarity, and a U-shaped function of π_2 when the cues were highly discriminable. Thus it is plain that $P(A_1 \mid T_1)$ does not become independent of π_2 with sufficient cue distinctiveness as predicted by the Burke–Estes model. Therefore, a test of the present theory will include its description of the dependence of $P(A_1 \mid T_1)$ on π_2 at different levels of cue similarity.

THEORY

The theory assumes that a subject has three strategies available in discriminative probability learning. The reinforcement contingencies determine what stategy is learned. That is, a subject is reinforced for performing according to one of the strategies, not for making a specific motor response. The strategies can be represented by three states of a homogeneous Markov chain.

CONDITIONING-STATE AXIOM:

On each trial the subject is in one of three possible states: appropriate (A), *unconditioned* (U), *inappropriate* (I). *It is assumed every subject starts the experiment in state* U *with probability one.*

RESPONSE AXIOM:

The response probability is dependent upon the state the subject is in. In state A, *the subject will make the appropriate response, which is the response most likely to be correct given the trial type, with probability one. In state* I, *the subject will make the inappropriate response (response most likely to be incorrect given the trial type) with probability one. In state* U, *the subject will make an A_1 response with probability p, regardless of the trial type.*

Notice that the subject in state U does not respond differentially on T_1 and T_2 trials.

The response probability p associated with state U can be assumed to grow according to a linear model and approach $P(E_1)$. Evidence for this assumption is available from studies of simple probability learning. However, there is nothing in the present theory which requires that p be equal to $P(E_1)$. For example, in certain cases, a subject in state U may overshoot $P(E_1)$. Hence, p can simply be estimated as a parameter of the model.

CONDITIONING AXIOMS:

C_1. *If the subject is in state* U *and makes an appropriate response and*

(a) *is correct, then with probability* c_1, *he goes into state* A *and with probability* $1 - c_1$, *he remains in state* U;

(b) *is incorrect, then with probability* c_2, *he goes into state* I *and with probability* $1 - c_2$, *he remains in state* U.

C_2. *If the subject is in state* U *and makes an inappropriate response and*

(a) *is correct, then with probability* c_1, *he goes into state I and with probability* $1 - c_1$, *he remains in state* U;

(b) *is incorrect, then with probability* c_2, *he goes into state* A *and with probability* $1 - c_2$, *he remains in state* U.

C_3. *If the subject is in state* A *and is incorrect, then with probability* c, *he goes into state* U *and with probability* $1 - c$, *he remains in state* A.

C_4. *If the subject is in state* I *and is incorrect, then with probability* c, *he goes into state* U *and with probability* $1 - c$, *he remains in state* I.

For the present development, it is assumed that $\pi_1 > .5$ and $\pi_2 < .5$. Therefore, appropriate responses are A_1 on T_1 trials and A_2 on T_2 trials. If $\pi_2 > .5$, an A_1 would be appropriate on T_2 trials. When $\pi_2 = .5$ no response can be identified as appropriate on T_2 trials. However, the experimenter can arbitrarily define A_1 (appropriate response on T_1 trials) as appropriate on T_2 trials and tests can be made of the model. By employing $\pi_1 > .5$ and $\pi_2 < .5$, we are developing the model for the traditional discrimination learning paradigm in which the subject learns to behave appropriately by responding differently on different trial types.

Figure 1 illustrates the transitions among conditioning states that are possible under the Conditioning and Response Axioms. For example, suppose that subject is in state U. A T_1 trial will occur with probability β. By the Response Axiom, the subject will make an A_1 response with probability p which will be reinforced by an appropriate event with probability π_1. Therefore, with probability c_1, the subject goes into state A and with probability $1 - c_1$, he remains in state U. These transitions among conditioning states lead to the following transition matrix $\mathbf{P}$ and response probability vectors for appropriate responses, $P(A_1 \mid T_1)$ and $P(A_2 \mid T_2)$:

$$\begin{array}{c} \\ \mathbf{P} = \begin{array}{c} \mathrm{A} \\ \mathrm{U} \\ \mathrm{I} \end{array} \end{array} \begin{array}{c} \begin{array}{ccc} \mathrm{A} & \mathrm{U} & \mathrm{I} \end{array} \\ \begin{bmatrix} 1 - c(1 - \pi_a) & c(1 - \pi_a) & 0 \\ E & 1 - E - F & F \\ 0 & \pi_a c & 1 - \pi_a c \end{bmatrix} \end{array} \begin{array}{c} P(A_1 \mid T_1) \\ \begin{bmatrix} 1 \\ p \\ 0 \end{bmatrix} \end{array} \begin{array}{c} P(A_2 \mid T_2) \\ \begin{bmatrix} 1 \\ 1 - p \\ 0 \end{bmatrix} \end{array}, \qquad (1)$$

where

$$\pi_a = \beta\pi_1 + (1 - \beta)(1 - \pi_2),$$
$$E = \beta\pi_1[pc_1 + (1 - p)c_2] + (1 - \beta)(1 - \pi_2)[pc_2 + (1 - p)c_1],$$

and

$$F = \beta(1 - \pi_1)[pc_2 + (1 - p)c_1] + (1 - \beta)\,\pi_2[pc_1 + (1 - p)c_2].$$

Asymptotic Statistics. For the present derivations, the states will be numbered as follows: 1 = A, 2 = U, and 3 = I. To derive the asymptotic probability of a particular conditioning state, let p_{ij} be the Markov chain represented by the matrix $\mathbf{P}$

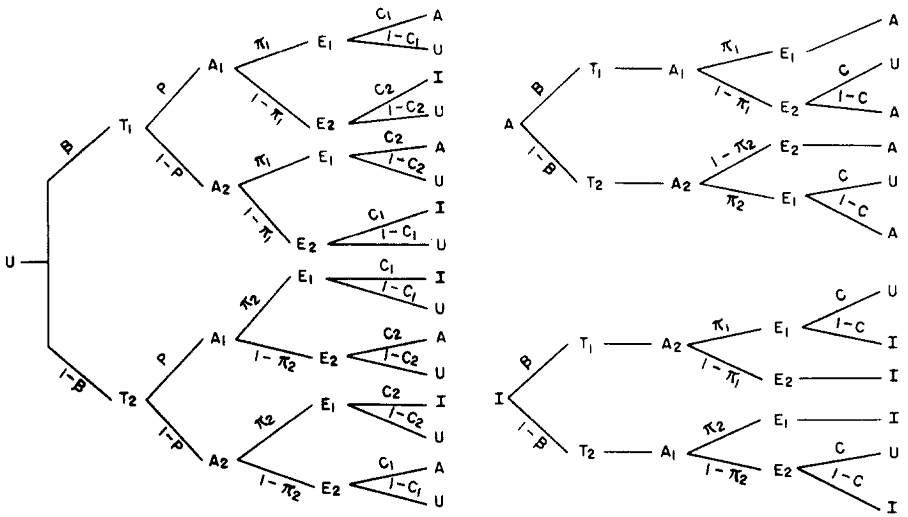

FIG. 1. Branching process, starting from each state on trial n, for the model when $\pi_1 > .5$ and $\pi_2 < .5$.

of Eq. 1, and define $p_{ij}^{(n)}$ as the probability of being in state j on trial $r + n$, given that at trial r the subject was in state i. The quantity is defined recursively:

$$p_{ij}^{(1)} = p_{ij}\,, \tag{2}$$

$$p_{ij}^{(n+1)} = \sum_v p_{iv} p_{vj}^{(n)}. \tag{3}$$

Moreover, if the limit exists and is independent of i, we set

$$u_j = \lim_{n\to\infty} p_{ij}^{(n)}. \tag{4}$$

The limiting quantities of u_j do exist for this finite-state Markov chain since it is irreducible and aperiodic. There are three states and at asymptote the probability of being in state i on trial $n+1$ is equal to the probability of being in state i on trial n ($p_{n+1} = p_n$). Let p_{ij} ($i, j = 1, 2, 3$) be the 3×3 transition matrix. We seek the values u such that $u_j = \sum_v u_v p_{vj}$ and $\Sigma u_j = 1$. The u_j's may be computed by the application of Cramér's rule:

$$u_j = \frac{D_j}{D_1 + D_2 + D_3}, \tag{5}$$

where $D_1 = E\pi_a$, $D_2 = c\pi_a(1 - \pi_a)$, and $D_3 = F(1 - \pi_a)$.

At asymptote, the probability of an A_1 response is a simple function of the u_j's. Redefining the states as A = 1, U = 2, and I = 3:

$$\lim_{n\to\infty} P(A_{1,n} \mid T_{1,n}) = A + pU$$
$$= \frac{E\pi_a + pc\pi_a(1 - \pi_a)}{E\pi_a + c\pi_a(1 - \pi_a) + F(1 - \pi_a)}, \tag{6}$$

and

$$\lim_{n\to\infty} P(A_{1,n} \mid T_{2,n}) = 1 - [A + (1 - p)\, U]$$
$$= 1 - \frac{E\pi_a + (1 - p)\, c\pi_a(1 - \pi_a)}{E\pi_a + c\pi_a(1 - \pi_a) + F(1 - \pi_a)}. \tag{7}$$

To simplify notation, the trial subscripts will be omitted when asymptotic expressions are referred to.

In some discrimination problems (e.g., concept identification) it has been shown that subjects abandon a strategy or hypothesis only if it leads to an incorrect response (Bower and Trabasso, 1964). If the present theory assumes that subjects change strategies only on error trials, the parameter c_1 will equal zero. Evidence for the validity of this assumption in the present learning task will be presented later. Also, to simplify the following derivations, it is assumed that $\pi_1 = 1 - \pi_2 = \pi_a$ and $\beta = 1 - \beta$. It follows from the Response Axiom that $p = P(E_1) = .5$. Therefore Eq. 6 and Eq. 7 reduce to:

$$P(A_1 \mid T_1) = \frac{.5\pi_a^2 + .5\pi_a(1 - \pi_a)\phi}{.5\pi_a^2 + \pi_a(1 - \pi_a)\phi + .5(1 - \pi_a)^2}, \tag{8}$$

$$P(A_1 \mid T_2) = 1 - P(A_1 \mid T_1), \tag{9}$$

$$\text{where } \phi = \frac{C}{C_2}.$$

The ratio ϕ can be taken as an index of differential responding on T_1 and T_2 trials.

That is, when ϕ becomes large, $P(A_1 \mid T_1) - P(A_1 \mid T_2)$ approaches zero. Whence by Eq. 8 and Eq. 9, respectively,

$$\lim_{\phi \to \infty} P(A_1 \mid T_1) = .5, \tag{10}$$

$$\lim_{\phi \to \infty} P(A_1 \mid T_2) = .5. \tag{11}$$

On the other hand, when ϕ becomes small, subjects will spend less time in state U relative to states A and I and respond differentially on T_1 and T_2 trials. It follows from Eqs. 8 and 9, respectively, that

$$\lim_{\phi \to 0} P(A_1 \mid T_1) = \frac{\pi_a^2}{\pi_a^2 + (1 - \pi_a)^2}, \tag{12}$$

$$\lim_{\phi \to 0} P(A_1 \mid T_2) = 1 - \frac{\pi_a^2}{\pi_a^2 + (1 - \pi_a)^2}. \tag{13}$$

Equations 10–13 show that discrimination learning increases with decreases in ϕ.

Sequential Statistics. The derivation of the first order conditional statistics is also straightforward. For example, given an A_1 on a T_1 trial, the subject was either in state A or U. If the subject was in state A, he remained there since an E_1 occurred; he therefore will make an A_1 on the present trial with probability one. If the subject was in state U (he made an A_1 with probability p), he can transit into state A with probability c_1, in which case he will make an A_1 on the present trial with probability one. If he remains in U, he will make an A_1 with probability p. The sum of these possible transition probabilities are divided by the sum of probabilities of being in the respective states. Thus, dropping the trial subscripts

$$\begin{aligned} P(A_{1,n+1} \mid T_{1,n+1}T_{1,n}A_{1,n}E_{1,n}) &= P(A_1 \mid T_1T_1A_1E_1) \\ &= \frac{\mathrm{A} + p\mathrm{U}[c_1 + (1 - c_1)p]}{\mathrm{A} + p\mathrm{U}}. \end{aligned} \tag{14}$$

The 16 independent first order conditionals are presented in the Appendix.

TESTS OF THE MODEL

In the estimation of the parameters, c_1 was found to be equal to or close to zero. Therefore, it seemed reasonable to assume that a subject could leave state U only if he was incorrect (i.e., $c_1 = 0$). This fact is appealing since the states represent different strategies and it seems likely that subjects will only change strategies when they are incorrect. In addition, p was assumed to be equal to $P(E_1) = \beta\pi_1 + (1 - \beta)\pi_2$. The two parameters, c and c_2, were estimated using a minimum χ^2 criterion between predicted and observed values for the 16 independent first order conditional probabilities. The data were pooled over all trials to increase the total number of observations. The predicted values were obtained by letting the probability vector at the start of trial 1 be

$$w_1 = [0 \quad 1 \quad 0]$$

and repeatedly computing the vector w_n for every trial n by the equation

$$w_n = w_{n-1}\mathbf{P}$$

where $\mathbf{P}$ is the transition matrix defined by Eq. 1. Then the entries of the average probability vector were taken as the values for the corresponding probabilities of the three states in the expressions for the conditionals presented in the Appendix.

Variations of π_1

Three values of π_1 were chosen while holding π_2 constant. Equation 1 shows that for given values of c, c_2, and π_2, the probability of being in state A is positively related to π_1. Therefore, increases in π_1 should increase the probability of an appropriate response.

METHOD

The subjects were 60 University of Massachusetts undergraduates assigned randomly to the experimental treatments. Up to four subjects were run at a time, each seated at a table top enclosure containing an Estes–Straughan conditioning board consisting of two white center cue lights labeled "loud" and "soft" and two red event lights positioned above each of two spring loaded lever switches. Tones were generated by a Hewlett-Packard model 200 audio oscillator and were presented over matched headphones with a continuous white masking noise of 70 db SPL. Experimental events were controlled by Lehigh Valley 1652 probability randomizers, Hunter interval timers, and relays. Events and responses were recorded on an Esterline-Angus event recorder.

The simultaneous onset of a tone and cue light started a trial. The two 800 Hz tones were 73 and 79 db SPL for the soft and loud tone, respectively.

The subjects were told that the cue light indicated whether the tone was a loud or soft tone and they were to predict which of the two red event lights would come on. The subjects had 2 sec. to make their prediction. At the offset of the tone, one of the event lights was illuminated indicating the correct prediction on that trial.

The intertrial interval was 7 sec. and each subject received 300 trials. The three groups differed by the schedules placed on the probability randomizers. The settings were $\beta = .5$, $\pi_2 = .2$, and $\pi_1 = .95$, .8, and .65 for Groups 1, 2, and 3, respectively. The exact stimulus and event schedules were not found to be significantly different from the a priori schedules.

RESULTS

Tables 1, 2, and 3 present the observed and predicted frequencies and proportions for Groups 1, 2, and 3, respectively. The significant χ^2's are to be expected from the large number of observations (Grant, 1962). For this reason and because of the fact that the frequencies contributing to each conditional differed widely, a better index of the description of the model was taken to be a weighted mean deviation (w.m.d.) between predicted and observed values. This statistic was computed for each group by multiplying the absolute deviation between predicted and observed values of each conditional by its joint frequency and averaging the 16 deviations. Tables 1, 2, and 3 show that the present model does very well in predicting the correct

TABLE 1

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR GROUP 1[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	1092	1065.9	.923	.901
1	1	1	2	57	54.2	.792	.753
1	1	2	1	104	87.5	.732	.616
1	1	2	2	8	5.1	.800	.505
1	2	1	1	43	30.5	.729	.516
1	2	1	2	165	127.5	.809	.625
1	2	2	1	220	195.3	.866	.769
1	2	2	2	897	880.7	.935	.918
2	1	1	1	181	156.2	.146	.126
2	1	1	2	16	25.6	.188	.301
2	1	2	1	49	61.8	.374	.472
2	1	2	2	5	10.9	.278	.603
2	2	1	1	29	40.4	.427	.595
2	2	1	2	69	92.8	.345	.464
2	2	2	1	60	69.5	.244	.283
2	2	2	2	113	111.7	.105	.103
1						.897	.849
2						.171	.186

[a] $\hat{c}_2 = .23$; $\hat{c} = .35$; $\pi_1 = .95$; $\chi^2 = 49.25$; and w.m.d. $= .040$.

TABLE 2

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR GROUP 2[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	871	880.3	.891	.901
1	1	1	2	161	153.5	.745	.711
1	1	2	1	95	88.9	.625	.585
1	1	2	2	31	19.8	.608	.388
1	2	1	1	28	18.4	.583	.383
1	2	1	2	134	102.0	.761	.580
1	2	2	1	194	165.1	.829	.706
1	2	2	2	925	882.4	.939	.896
2	1	1	1	91	92.4	.092	.093
2	1	1	2	48	66.6	.200	.276
2	1	2	1	36	58.4	.245	.397
2	1	2	2	30	31.7	.556	.587
2	2	1	1	19	34.9	.322	.592
2	2	1	2	40	68.9	.234	.403
2	2	2	1	55	76.0	.204	.281
2	2	2	2	133	114.6	.114	.098
1						.859	.814
2						.146	.175

[a] $\hat{c}_2 = .39$; $\hat{c} = .37$; $\pi_1 = .8$; $\chi^2 = 71.34$; and w.m.d. $= .046$.

rank-orderings of the conditionals. In comparison, the reinforcement models, such as the Burke–Estes and Atkinson models, cannot predict the rank-ordering when $T_{i,n+1} \neq T_{i,n}$. Across the three groups, only three inversions are observed for the present Markov model while 23 inversions exist for the predictions of reinforcement models.

Even though the present model provides a good description of the data, the parameter values should also be consistent with the assumptions underlying the psychological processes. As shown in the tables, the parameter c is fairly invariant (.35 to .39) as is c_2 with the exception of Group 1. The invariance of the parameters in these three groups is, of course, predicted by the assumptions of the model.

Variations of β

In studies where $\pi_1 = 1 - \pi_2$, variations of β only affect the absolute event probabilities. Since $p = P(E_1)$ is a positive function of the probability of T_1 trials, the response probability p associated with state U increases with increases in β. Thus more overshooting on T_1 trials is predicted for increases in β for given values of c and c_2.

METHOD

The method has been reported elsewhere (Myers and Cruse, 1968). Three groups of 20 subjects each were given 500 trials. The π_1 was .85, π_2 was .15, and the β values were either .75, .50, or .25 for the three groups. For the subsequent analysis, T_1 trials for the $\beta = .75$ condition

TABLE 3

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR GROUP 3[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	560	593.5	.818	.867
1	1	1	2	245	249.0	.637	.649
1	1	2	1	159	145.1	.589	.537
1	1	2	2	63	55.3	.358	.314
1	2	1	1	17	10.2	.472	.284
1	2	1	2	133	118.0	.564	.500
1	2	2	1	124	106.2	.709	.607
1	2	2	2	828	844.0	.803	.819
2	1	1	1	63	62.0	.090	.088
2	1	1	2	59	92.1	.168	.263
2	1	2	1	79	90.5	.305	.349
2	1	2	2	55	84.4	.342	.524
2	2	1	1	18	27.4	.375	.571
2	2	1	2	46	70.1	.256	.390
2	2	2	1	54	59.6	.273	.301
2	2	2	2	122	128.5	.114	.120
1						.711	.709
2						.167	.207

[a] $\hat{c}_2 = .40$; $\hat{c} = .39$; $\pi_1 = .65$; $\chi^2 = 50.09$; and w.m.d. = .039.

TABLE 4

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR $\beta = .75$[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	9642	9386.7	.973	.947
1	1	1	2	1526	1428.7	.842	.788
1	1	2	1	695	612.9	.743	.656
1	1	2	2	96	78.0	.561	.456
1	2	1	1	52	38.9	.732	.548
1	2	1	2	340	294.8	.852	.739
1	2	2	1	326	316.6	.858	.833
1	2	2	2	2147	2235.5	.934	.972
2	1	1	1	288	277.0	.115	.111
2	1	1	2	75	130.0	.198	.343
2	1	2	1	55	106.7	.254	.492
2	1	2	2	24	39.8	.471	.781
2	2	1	1	11	12.5	.647	.736
2	2	1	2	42	39.6	.438	.413
2	2	2	1	25	39.3	.188	.296
2	2	2	2	50	34.0	.086	.058
1						.927	.900
2						.143	.171

[a] $\hat{c}_2 = .55$; $\hat{c} = .36$; $\chi^2 = 112.01$; and w.m.d. $= .038$.

were pooled with T_2 trials for $\beta = .25$ to increase the low number of observations on the infrequent trial types. Therefore, 40 subjects are in the $\beta = .75$ group while 20 subjects are in the $\beta = .5$ group.

RESULTS

Tables 4 and 5 show that the model accurately describes manipulations of β. The model predicts a larger spread in each set of four conditionals than that observed in the data. Nevertheless, the model does predict the overshooting found on T_1 trials when $\beta = .75$ which cannot be predicted by reinforcement models. The fact that c_2 of the $\beta = .75$ condition is about three times c_2 of the $\beta = .5$ condition indicates that subjects are more likely to leave state U when there is a predominance of one trial type. Since c did not differ under the two conditions, the larger value of c_2 and hence the lower value of ϕ can account for the better discrimination, $P(A_1 \mid T_1) - P(A_1 \mid T_2)$, under the $\beta = .75$ condition.

Cue Similarity

In the previous analyses, the T_1 and T_2 trials were highly discriminable. The present theory predicts that confusable T_1 and T_2 trials will make it difficult for a subject to test the strategies reliably. A subject must identify the trial type before he can perform according to an appropriate or inappropriate strategy. Therefore, if a subject does not observe the trial type, he will not respond differentially on the different trial types. When the subject does not observe the trial type, he is in state U and will

TABLE 5

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR $\beta = .5$[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	1329	1361.0	.843	.863
1	1	1	2	180	173.9	.738	.713
1	1	2	1	277	238.3	.643	.553
1	1	2	2	24	25.8	.429	.461
1	2	1	1	50	36.4	.633	.461
1	2	1	2	288	257.7	.618	.553
1	2	2	1	256	220.9	.826	.713
1	2	2	2	1546	1538.7	.867	.863
2	1	1	1	258	238.2	.148	.137
2	1	1	2	102	114.9	.255	.287
2	1	2	1	165	192.2	.384	.447
2	1	2	2	40	48.5	.444	.539
2	2	1	1	42	40.4	.560	.539
2	2	1	2	220	237.4	.414	.447
2	2	2	1	80	84.8	.271	.287
2	2	2	2	212	209.5	.139	.137
1						.799	.779
2						.220	.229

[a] $\hat{c}_2 = .17$; $\hat{c} = .35$; $\chi^2 = 31.75$; and w.m.d. $= .025$.

simply match $P(E_1)$. It follows that the parameters c and c_2 should increase and decrease respectively with increases in cue similarity. A test of this prediction was carried out at three levels of cue similarity with an independent observation of the confusability of the stimuli from half the subjects.

METHOD

The method has been reported elsewhere (Massaro, *et al.*, 1968). Three levels of cue similarity (tones differing by 1.5, 3, and 6 db) were employed at a $\pi_1 = .8$ and $\pi_2 = .2$ reinforcement condition. Half the subjects were also required to indicate which of the two tones (loud or soft) was being presented. These subjects were given no information regarding their identification response. The mean percentage of correct identifications for the 1.5, 3, and 6 db conditions were 63, 79, and 93 %, respectively. Since the identification response had no significant effect on event predictions, these data were pooled at each level of cue similarity. Therefore, each group had 24 subjects with 300 observations per subject.

TABLE 6

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR HIGHLY SIMILAR CUES (1.5 db differential)[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	534	539.0	.642	.648
1	1	1	2	92	100.7	.482	.527
1	1	2	1	288	218.2	.699	.530
1	1	2	2	77	75.5	.497	.487
1	2	1	1	105	86.2	.593	.487
1	2	1	2	307	313.5	.519	.530
1	2	2	1	174	134.4	.682	.527
1	2	2	2	582	580.5	.650	.648
2	1	1	1	300	315.5	.335	.352
2	1	1	2	79	114.4	.326	.473
2	1	2	1	298	242.8	.578	.471
2	1	2	2	69	72.8	.486	.513
2	2	1	1	61	63.6	.492	.513
2	2	1	2	240	211.2	.534	.471
2	2	2	1	71	77.6	.433	.473
2	2	2	2	312	296.9	.370	.352
1						.615	.583
2						.424	.413

[a] $\hat{c}_2 = .07$; $\hat{c} = .65$; $\chi^2 = 68.94$; and w.m.d. $= .045$.

RESULTS

Tables 6, 7, and 8 indicate that the present model does as well with similar cues as in the previously described experiments with distinctive cues. The fact that c_2 increased (from .07 to .19 to .27) and c decreased (from .65 to .53 to .45) with increases in cue distinctiveness indicates that the parameters of the model change as predicted from the psychological assumptions. It is interesting that given a high level of cue similarity (cf. Table 6), the spread of each set of four conditionals is not as large as it was for the distinctive cue groups. This decrease in spread is predicted nicely by the Markov model, even though the observed rank-orderings are not as orderly as those observed with distinctive cues. On the other hand, the reinforcement models would predict an even larger spread among conditionals when the cues are more similar when $T_{i,n+1} \neq T_{i,n}$. The Markov model also gives a better description of the reinforcement effects at different levels of cue similarity.

TABLE 7

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR AN INTERMEDIATE LEVEL OF CUE SIMILARITY (3.0 db DIFFERENTIAL)[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	886	820.8	.854	.791
1	1	1	2	172	158.2	.649	.597
1	1	2	1	203	171.6	.672	.568
1	1	2	2	40	32.6	.563	.459
1	2	1	1	94	58.3	.740	.459
1	2	1	2	260	225.5	.655	.568
1	2	2	1	236	180.3	.781	.597
1	2	2	2	888	858.7	.818	.791
2	1	1	1	238	245.0	.203	.209
2	1	1	2	59	118.5	.201	.403
2	1	2	1	124	145.5	.368	.432
2	1	2	2	28	48.1	.315	.541
2	2	1	1	43	40.6	.573	.541
2	2	1	2	132	144.3	.395	.432
2	2	2	1	69	85.1	.327	.403
2	2	2	2	226	221.2	.214	.209
1						.775	.698
2						.258	.294

[a] $\hat{c}_2 = .19$; $\hat{c} = .53$; $\chi^2 = 105.16$; and w.m.d. = .058.

TABLE 8

OBSERVED (OBS.) AND PREDICTED (PRED.) VALUES OF $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$ AND $P(A_{1,n+1} \mid T_{i,n+1})$ FOR HIGHLY DISTINCTIVE CUES (6.0 db DIFFERENTIAL)[a]

				Frequency		Proportion	
$T_{i,n+1}$	$T_{j,n}$	$A_{k,n}$	$E_{\ell,n}$	Obs.	Pred.	Obs.	Pred.
1	1	1	1	924	898.3	.873	.849
1	1	1	2	179	171.3	.681	.651
1	1	2	1	186	164.7	.657	.582
1	1	2	2	30	34.8	.375	.436
1	2	1	1	68	47.5	.624	.436
1	2	1	2	247	211.9	.679	.582
1	2	2	1	222	211.0	.685	.651
1	2	2	2	983	951.0	.878	.849
2	1	1	1	196	179.9	.164	.151
2	1	1	2	78	106.0	.257	.349
2	1	2	1	122	130.8	.390	.418
2	1	2	2	36	47.4	.429	.565
2	2	1	1	40	32.7	.690	.565
2	2	1	2	133	144.6	.384	.418
2	2	2	1	83	79.5	.364	.349
2	2	2	2	177	157.3	.170	.151
1						.788	.747
2						.243	.246

[a] $\hat{c}_2 = .27$; $\hat{c} = .45$; $\chi^2 = 38.20$; and w.m.d. $= .036$.

LIMITATIONS OF THE MODEL

To some extent, the deviations of the present model from the data can be discounted as arising from unreliable variations in the data. However, the larger spread in the predicted conditionals than the observed may be due to the definition of states in the present model. That is, it was postulated that the change of states from one trial to the next is Markovian. However, several investigators (e.g., Jones and Myers, 1966) have shown that subjects do respond to run lengths in probability learning. Therefore, transitions between states may depend upon more than the previous trial and this variance will not be accounted for by the present model.

The present model, like most Markov models, has assumed that the transition probabilities between states are invariant. This assumption of homogeneity, although a mathematically tractable property, probably only approximates the learning process. That is, early in training, transitions out of state U should not necessarily be as likely as later in the learning process.

A FOUR STATE MODEL

In discriminative probability learning a subject may not observe the trial type and, therefore, not respond differentially on different trial types. However, even if the subject observes the trial type, he will not necessarily be conditioned to predict the more or less frequent event given the trial type. In the present model the two above conditions are grouped into state U. An extension of the model would treat each of these as different states.

The four state model would have a state where the subject does not observe the trial type, $\bar{O}$. When the subject observes the trial type, he is in one of the states of the three state model. The probability of observing would be expected to increase over training reaching an asymptote positively related to cue distinctiveness and π_a. The response vectors for $\bar{O}$ and U could now differ. In state U, the subject may match the event probability given the trial type since he has observed the trial type. Whereas in state $\bar{O}$, the subject can either match $P(E_1)$ or respond at a .5 chance level.

CONCLUSION

The Markov model presented in this paper has shown to predict, at least qualitatively, observed data from discrimination learning studies. The psychological assumption that discrimination learning is a form of stimulus learning rather than response learning has been able to account for overshooting results and reinforcements effects that previous models have been unable to describe. The cue similarity studies also support the model since the parameter values change as would be predicted from the psychological assumptions. Another attraction of the present model is the parsimony of the mathematical development. That is, the present model is mathematically simpler than both the Burke–Estes and the Atkinson models.

The suggested modifications of the present model or other theoretical approaches may improve the description of two-stimulus probability learning. Nevertheless, it is clear that the proposed three state model provides the best description presently available.

APPENDIX

Presented below are the expressions for the conditional statistics of the form $P(A_{1,n+1} \mid T_{i,n+1}T_{j,n}A_{k,n}E_{\ell,n})$

$$P(A_1 \mid T_1T_1A_1E_1) = \frac{\mathrm{A} + p\mathrm{U}[c_1 + (1 - c_1)p]}{\mathrm{A} + p\mathrm{U}}, \tag{14}$$

$$P(A_1 \mid T_1T_1A_1E_2) = \frac{\mathrm{A}[cp + (1 - c)] + \mathrm{U}(1 - c_2)p^2}{\mathrm{A} + p\mathrm{U}}, \tag{15}$$

$$P(A_1 \mid T_1T_1A_2E_1) = \frac{\mathrm{I}cp + (1 - p)\,\mathrm{U}[c_2 + (1 - c_2)p]}{\mathrm{I} + (1 - p)\mathrm{U}}, \tag{16}$$

$$P(A_1 \mid T_1T_1A_2E_2) = \frac{(1 - p)\,\mathrm{U}(1 - c_1)p}{\mathrm{I} + (1 - p)\mathrm{U}}, \tag{17}$$

$$P(A_1 \mid T_1T_2A_1E_1) = \frac{\mathrm{U}(1 - c_1)p^2}{\mathrm{I} + p\mathrm{U}}, \tag{18}$$

$$P(A_1 \mid T_1T_2A_1E_2) = \frac{\mathrm{I}cp + p\mathrm{U}[c_2 + (1 - c_2)p]}{\mathrm{I} + p\mathrm{U}}, \tag{19}$$

$$P(A_1 \mid T_1T_2A_2E_1) = \frac{\mathrm{A}[cp + (1 - c)] + (1 - p)\,\mathrm{U}(1 - c_2)p}{\mathrm{A} + (1 - p)\mathrm{U}}, \tag{20}$$

$$P(A_1 \mid T_1T_2A_2E_2) = \frac{\mathrm{A} + (1 - p)\,\mathrm{U}[c_1 + (1 - c_1)p]}{\mathrm{A} + (1 - p)\mathrm{U}}, \tag{21}$$

$$P(A_1 \mid T_2T_1A_1E_1) = \frac{\mathrm{U}(1 - c_1)p^2}{\mathrm{A} + p\mathrm{U}}, \tag{22}$$

$$P(A_1 \mid T_2T_1A_1E_2) = \frac{\mathrm{A}cp + p\mathrm{U}[c_2 + (1 - c_2)p]}{\mathrm{A} + p\mathrm{U}}, \tag{23}$$

$$P(A_1 \mid T_2T_1A_2E_1) = \frac{\mathrm{I}[cp + (1 - c)] + (1 - p)\,\mathrm{U}(1 - c_2)p}{\mathrm{I} + (1 - p)\mathrm{U}}, \tag{24}$$

$$P(A_1 \mid T_2T_1A_2E_2) = \frac{\mathrm{I} + (1 - p)\,\mathrm{U}[c_1 + (1 - c_1)p]}{\mathrm{I} + (1 - p)\mathrm{U}}, \tag{25}$$

$$P(A_1 \mid T_2T_2A_1E_1) = \frac{\mathrm{I} + p\mathrm{U}[c_1 + (1 - c_1)p]}{\mathrm{I} + p\mathrm{U}}, \tag{26}$$

$$P(A_1 \mid T_2T_2A_1E_2) = \frac{\mathrm{I}[cp + (1 - c)] + \mathrm{U}(1 - c_2)p^2}{\mathrm{I} + p\mathrm{U}}, \tag{27}$$

$$P(A_1 \mid T_2T_2A_2E_1) = \frac{\mathrm{A}cp + (1 - p)\,\mathrm{U}[c_2 + (1 - c_2)p]}{\mathrm{A} + (1 - p)\mathrm{U}}, \tag{28}$$

$$P(A_1 \mid T_2T_2A_2E_2) = \frac{(1 - p)\,\mathrm{U}(1 - c_1)p}{\mathrm{A} + (1 - p)\mathrm{U}}. \tag{29}$$

REFERENCES

Atkinson, R. C. A Markov model for discrimination learning. *Psychometrika*, 1958, **23**, 309–322.

Bower, G. H., and Trabasso, T. R. Concept identification. *In* R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford: Stanford Univer. Press, 1964. Pp. 32–94.

Burke, C. J., and Estes, W. K. A component model for stimulus variables in discrimination learning. *Psychometrika*, 1957, **22**, 133–145.

Grant, D. A. Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychological Review*, 1962, **69**, 54–61.

Jones, M. R., and Myers, J. L. A comparison to two methods of event randomization in probability learning. *Journal of Experimental Psychology*, 1966, **72**, 909–911.

Lee, W. Conditioning parameter model for reinforcement generalization in probabilistic discrimination learning. *Journal of Mathematical Psychology*, 1966, **3**, 184–196.

Massaro, D. W., Halpern, J., and Moore, J. W. Generalization effects in human discrimination learning with overt cue identification. *Journal of Experimental Psychology*, 1968, **77**, 474–482.

Moore, J. W., and Halpern, J. Two-choice discrimination learning as a function of stimulus similarity along an auditory intensity dimension. *Psychonomic Science*, 1965, **3**, 441–442.

Myers, J. L., and Cruse, D. Two-choice discrimination learning as a function of stimulus and event probabilities. *Journal of Experimental Psychology*, 1968, **77**, 453–459.

Popper, J., and Atkinson, R. C. Discrimination learning in a verbal conditioning situation. *Journal of Experimental Psychology*, 1958, **56**, 21–25.

Received: July 31, 1967