Reprinted from STUDIES IN MATHEMATICAL LEARNING THEORY
Robert R. Bush and William K. Estes, Eds.
Published by Stanford University Press, Stanford, California

3

Applications of a Markov Model to Two-Person Noncooperative Games[1]

RICHARD C. ATKINSON, University of California at Los Angeles
PATRICK SUPPES, Stanford University

The present study represents an attempt to account for quantitative properties of behavior in a game situation involving *social interaction* between two individuals. The basis of prediction is a Markov model for learning which, in conceptual development, is closely related to statistical learning theory [5].

Before proceeding to the details of this study there are three general remarks we would like to make concerning the fundamental ideas and methods:

(1) The principles of behavior that constitute our theory of social interaction are rigorously derivable from general principles of individual behavior, and in particular from stochastic versions of reinforcement theory.

(2) The results are quantitative in a sharp sense; elaborate mathematics have not been applied to quantities which can only be ordinally measured.

(3) The underlying principles constitute a genuine theory in the sense that prior to experimentation quantitative predictions of behavior may be made for a wide range of parameter values.

For the purposes of this experiment a play of a game is a trial. On a given trial, each of the players makes a choice between two responses. After the players have independently indicated their responses, the outcome of the trial is announced. In an earlier study [3] we considered games where the outcomes were such that on each trial one player was "correct" and the other player was "incorrect"; that is, zero-sum games. In this study the games have outcomes such that on each trial both players can be correct, both can be incorrect, or one can be correct and the other incorrect; that is, they are non-zero-sum games. More important than the shift from zero-sum to non-zero-sum games is the fact that in this experiment, in contrast to [3], the players are informed that they are interacting with each other.

On all trials, the game is described by the following payoff matrix:

[1] This research was supported by the Behavioral Sciences Division of the Ford Foundation and by the Group Psychology Branch of the Office of Naval Research.

	B_1	B_2
A_1	(a_1, b_1, c_1, d_1)	(a_2, b_2, c_2, d_2)
A_2	(a_3, b_3, c_3, d_3)	(a_4, b_4, c_4, d_4)

The players are designated A and B. The responses available to A are A_1 and A_2; similarly, the responses for B are B_1 and B_2. If A selects A_1 and B selects B_1, then there is (i) a probability a_1 that both players are correct, (ii) a probability b_1 that A is correct and B is incorrect, (iii) a probability c_1 that A is incorrect and B is correct, and (iv) a probability d_1 that both players are incorrect. These outcomes are mutually exclusive and exhaustive, given the response pair A_1B_1. The outcomes of the other three response pairs, A_1B_2, A_2B_1, and A_2B_2, are similarly specified in terms of (a_2, b_2, c_2, d_2), (a_3, b_3, c_3, d_3) and (a_4, b_4, c_4, d_4).

The experiment to be presented employs the procedure outlined above. At the start of an experimental session, the subjects are informed of the game characteristics of the situation and instructed to maximize the number of trials on which their responses are correct.

Model

Since a detailed mathematical development of the model will be presented elsewhere, we shall confine ourselves to the most salient features and omit mathematical proofs. A more complete development of the psychological concepts which lead to the present model and a consideration of its relation to the Estes and Burke stimulus-sampling theory can be found in Atkinson and Suppes [3].

We begin by making the simplifying assumption that on all trials a player's response is determined by a single stimulus—that is, the event associated with the onset of a trial. The subject is described as being in one of two possible states: (a) if he is in State 1, the stimulus is conditioned to Response 1, and in the presence of the stimulus Response 1 will be elicited; (b) if he is in State 2, the stimulus is conditioned to Response 2, and in the presence of the stimulus Response 2 will be elicited. Thus, on any trial n, the two players are described in terms of one of the four states $\langle A_1, B_1\rangle$, $\langle A_1, B_2\rangle$, $\langle A_2, B_1\rangle$, and $\langle A_2, B_2\rangle$, where the first member of a couple indicates the state of Player A and the second, the state of Player B. For example, $\langle A_2, B_1\rangle$ means that Player A will make Response A_2 and Player B will make B_1. It is postulated that the change of states from one trial to the next is Markovian, and learning is defined with respect to the $\langle A_i, B_j\rangle$ states. This Markov property may be derived from more general independence-of-path assumptions, which are not usually explicitly stated but are characteristic of stimulus-sampling theory.

In specifying the learning process the term *reinforcement* will be used. It is defined as follows for a situation in which only two responses are available to the subject: if a response occurs and is correct, then the response is reinforced; if a response occurs and is incorrect, then the alternative response is reinforced. Thus on every trial one of the two responses is reinforced.

When one of Player A's responses is reinforced on trial n, there is (i) a probability θ_A that the stimulus governing Player A's response will be conditioned to the reinforced response and therefore that on trial $n+1$ Player A will make the response reinforced on trial n; and (ii) a probability $1-\theta_A$ that the conditioned status of the stimulus will remain unchanged and therefore that on trial $n+1$ Player A will repeat the response made on trial n. Identical rules describe the learning process for Player B in terms of θ_B.

The parameters θ_A and θ_B describe the learning rate characteristics of Players A and B, respectively. Some theoretical predictions can be made without knowledge of these parameter values; in general, however, predictions are a function of both θ_A and θ_B. These values can be estimated from a subset of the data and used to predict the remaining data ([4], [7]), or in some cases they can be estimated from other experiments ([2], [6]).

For the set of assumptions given above and the payoff probabilities a_i, b_i, c_i, and d_i, the transition matrix ([8], [9]) describing the learning process can be derived and is as follows:

$$P=\begin{bmatrix}
\begin{array}{l} a_1+b_1(1-\theta_B)\\ +c_1(1-\theta_A)\\ +d_1(1-\theta_A)(1-\theta_B)\end{array} & \begin{array}{l} b_1\theta_B\\ +d_1\theta_B(1-\theta_A)\end{array} & \begin{array}{l} c_1\theta_A\\ +d_1\theta_A(1-\theta_B)\end{array} & d_1\theta_A\theta_B \\[2ex]
\begin{array}{l} b_2\theta_B\\ +d_2\theta_B(1-\theta_A)\end{array} & \begin{array}{l} a_2+b_2(1-\theta_B)\\ +c_2(1-\theta_A)\\ +d_2(1-\theta_A)(1-\theta_B)\end{array} & d_2\theta_A\theta_B & \begin{array}{l} c_2\theta_A\\ +d_2\theta_A(1-\theta_B)\end{array} \\[2ex]
\begin{array}{l} c_3\theta_A\\ +d_3\theta_A(1-\theta_B)\end{array} & d_3\theta_A\theta_B & \begin{array}{l} a_3+b_3(1-\theta_B)\\ +c_3(1-\theta_A)\\ +d_3(1-\theta_A)(1-\theta_B)\end{array} & \begin{array}{l} b_3\theta_B\\ +d_3\theta_B(1-\theta_A)\end{array} \\[2ex]
d_4\theta_A\theta_B & \begin{array}{l} c_4\theta_A\\ +d_4\theta_A(1-\theta_B)\end{array} & \begin{array}{l} b_4\theta_B\\ +d_4\theta_B(1-\theta_A)\end{array} & \begin{array}{l} a_4+b_4(1-\theta_B)\\ +c_4(1-\theta_A)\\ +d_4(1-\theta_A)(1-\theta_B)\end{array}
\end{bmatrix}$$

In the matrix P the rows and columns correspond to the following states: (1) $\langle A_1, B_1\rangle$, (2) $\langle A_1, B_2\rangle$, (3) $\langle A_2, B_1\rangle$, and (4) $\langle A_2, B_2\rangle$. Rows designate the state on trial n and columns the state on trial $n+1$. Thus, for example, $p_{23}=d_2\theta_A\theta_B$ (the entry in row 2, column 3) is the conditional probability of being in State 3 on trial $n+1$ given that the pair of subjects was in State 2 on trial n, for we have

$$d_2\theta_A\theta_B = a_2\cdot 0 + b_2\cdot 0 + c_2\cdot 0 \\ + d_2[(1-\theta_A)(1-\theta_B)\cdot 0 + \theta_A(1-\theta_B)\cdot 0 + \theta_B(1-\theta_A)\cdot 0 + \theta_A\theta_B\cdot 1]\,.$$

Specifically, a transition from State 2 to State 3 can occur only if A_1 and B_2 both were incorrect (probability d_2) and conditioning occurred for both players (probability $\theta_A\cdot\theta_B$).

The transition probabilities p_{ij} and an initial probability distribution for the states completely describe behavior in the situation, and from these one can obtain any theoretical quantity desired. Of particular interest is an expression for the asymptotic probabilities of each of the four states, that is, the asymptotic distribution for the joint occurrence of responses A_i and $B_j (i, j = 1 \text{ or } 2)$. This quantity will be denoted as $u_k (k = 1, 2, 3, 4)$. It can be shown that the probability distribution $\{u_k\}$ satisfies the following system of equations:

$$(1) \qquad u_k = \sum_{i=1}^{4} u_i p_{ik} \qquad (k = 1, 2, 3, 4) .$$

In terms of the quantities u_k, we obtain the asymptotic probability of an A_1 and a B_1 response, namely,

$$(2) \qquad p_\infty(A_1) = u_1 + u_2 ,$$

$$(3) \qquad p_\infty(B_1) = u_1 + u_3 .$$

The general expressions for $p_\infty(A_1)$ and $p_\infty(B_1)$ are too lengthy to reproduce here, but special forms of them will be used in analyzing data of the present study.

The essential interactional character of our Markov model is made clear by the following observation. The probability of the joint occurrence of responses A_i and B_j on trial n, $p_n(A_i \cap B_j)$, is fundamental to the process, rather than the individual probabilities $p_n(A_i)$ and $p_n(B_j)$. In particular, the responses of the players are not independent; that is, in general

$$p_n(A_i \cap B_j) \neq p_n(A_i) \cdot p_n(B_j) .$$

Method

Experimental Parameter Values. Two groups were run; for game-theoretic considerations to be indicated later, they were designated Sure and Mixed. For the Sure Group:

$a_1 = 1/4$	$a_2 = 3/4$	$a_3 = 5/32$	$a_4 = 5/16$
$b_1 = 3/4$	$b_2 = 0$	$b_3 = 3/32$	$b_4 = 3/16$
$c_1 = 0$	$c_2 = 1/4$	$c_3 = 15/32$	$c_4 = 5/16$
$d_1 = 0$	$d_2 = 0$	$d_3 = 9/32$	$d_3 = 3/16$

For the Mixed Group:

$a_1 = 0$	$a_2 = 3/8$	$a_3 = 0$	$a_4 = 15/64$
$b_1 = 1$	$b_2 = 0$	$b_3 = 0$	$b_4 = 25/64$
$c_1 = 0$	$c_2 = 5/8$	$c_3 = 5/8$	$c_4 = 9/64$
$d_1 = 0$	$d_2 = 0$	$d_3 = 3/8$	$d_4 = 15/64$

Apparatus. The apparatus has already been described in detail [3], and only the salient features will be repeated here. The subjects, run in pairs, sat at opposite ends of a table Mounted vertically in front of each subject was a large opaque panel. The experimenter sat between the two panels and was not visible to either subject. The apparatus, as viewed from the subject's side, consisted of two silent operating keys mounted at the base of the panel; upon the panel were mounted three milk-glass panel lights. One of these lights, which served as the signal for the subject to respond, was centered between the keys at the subject's eye level. Each of the two remaining lights, the reinforcing signals, was mounted directly above one of the keys. The presentation and duration of the lights were automatically controlled.

Subjects. The subjects were 88 undergraduates obtained from introductory psychology courses. They were randomly assigned to the experimental groups with the restriction that there were 24 pairs of subjects in the Sure Group and 20 in the Mixed Group.

Procedure. For each pair of subjects, one was randomly selected as Player A and the other as Player B. Further, for each subject one of the two response keys was randomly designated Response 1 and the other Response 2, with the restriction that the following possible combinations occurred equally often in each of the experimental groups: (a) A_1 and B_1 on the right, (b) A_1 on the right and B_1 on the left, (*c*) A_1 on the left and B_1 on the right, and (d) A_1 and B_1 on the left.

When the subjects had been seated, they were read the following instructions:

"This experiment is analogous to a real-life situation where what you gain or lose depends not only on what you do but also on what someone else does. In fact, you should think of the situation as a game involving you and another player, the person at the other end of the table.

"The experiment for each of you consists of a series of trials. The top center lamp on your panel will light for about two seconds to indicate the start of each trial. Shortly thereafter one or the other of the two lower lamps will light up. Your job is to predict on each trial which one of the two lower lamps will light and indicate your prediction by pressing the proper key. That is, if you expect the left lamp to light, press the left key; if you expect the right lamp to light, press the right key. On each trial press one or the other of the two keys, but never both. If you are not sure which key to press, then guess.

"Be sure to indicate your choice by pressing the proper key immediately after the onset of the signal light. That is, when the signal light goes on, press one or the other key down and release it. Then wait until one of the lower lights goes on. If the light above the key you pressed goes on, your prediction was correct; if the light above the key opposite from the one you pressed goes on, you were incorrect.

"Being correct or incorrect on a given trial depends on the key you press and also on the key the other player presses. With some combinations of your key choice with the other player's key choice, you may both be correct; with other combinations one player will be correct and the other incorrect; for still other combinations you may both be incorrect.

"As you have probably already guessed, the situation is fairly complicated. The object of the experiment is to see how many correct predictions you can get over a series of trials."

Questions were answered by paraphrasing the appropriate part of the instructions.

Following the instructions, 210 trials were run in continuous sequence. For each pair of subjects, sequences of reinforcing lights were generated in accordance with assigned values of (a_i, b_i, c_i, d_i) and observed responses.

On all trials the signal light was lighted for 3.5 seconds; the time between successive signal exposures was 10 seconds. The reinforcing light followed the cessation of the signal light by 1.5 seconds and remained on for 2 seconds.

Results and Discussion

Mean Learning Curves and Asymptotic Results. Fig. 1 presents the mean proportions of A_1 and B_1 responses in successive blocks of 30 trials for the entire sequence of 210 trials. An inspection of this figure indicates that responses were fairly stable over the last 90 trials. To check the stability of response probabilities for individual data, t's for paired measures were computed between response proportions for the first and last halves of the final block of 90 trials. In all cases the obtained values of t did not approach significance at the .10 level. In view of these results it appears reasonable to assume that a constant level of responding has been attained; consequently, the proportions computed over the last 90 trials were used as estimates of

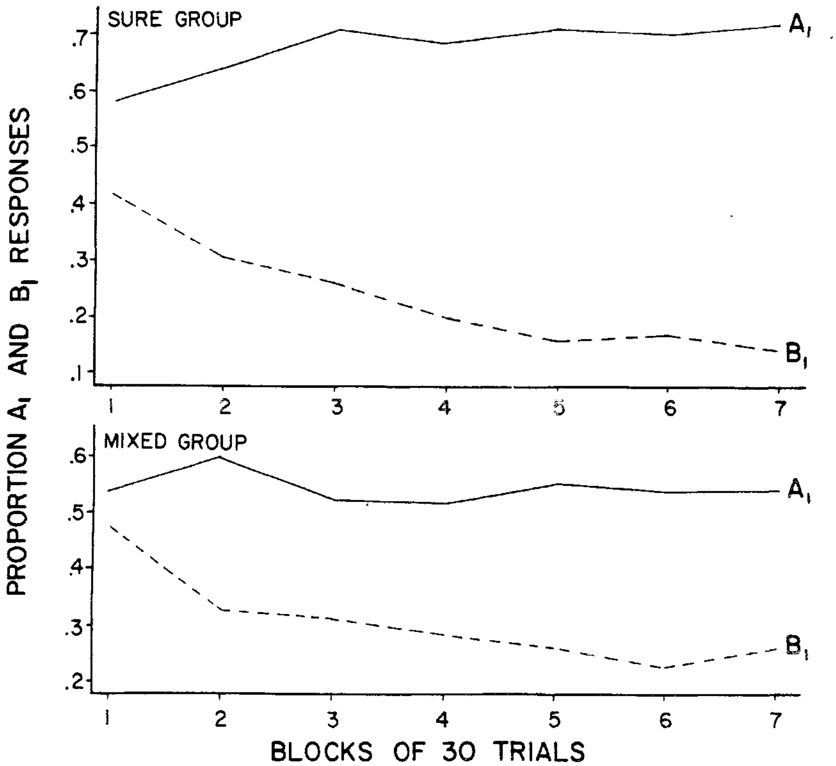

FIGURE 1. The mean proportions of A_1 and B_1 responses in successive blocks of 30 trials for the Sure and Mixed Groups.

the asymptotic probabilities of A_1 and B_1 responses. Table 1 presents the observed mean proportions of A_1 and B_1 responses in the last 90-trial block and the standard deviations associated with these means. Entries for the Sure Group are based on $N = 24$; for the Mixed Group, $N = 20$.

TABLE 1

Predicted and Observed Mean Proportions of A_1 and B_1 Responses over the Last 90 Trials

Group	A_1			B_1		
	Predicted	Observed	s	Predicted	Observed	s
Sure	.71	.72	.090	.14	.16	.099
Mixed	.55	.55	.056	.28	.25	.079

The values predicted by the Markov model are also presented in Table 1. They are obtained by substitution in the following equations:

$$(4)\quad p_\infty(A_1) = \frac{(b_3+d_3)(c_4+d_4)+(c_3+d_3)(b_4+d_4)}{(c_2+d_2+c_4+d_4)(b_3+d_3+b_4+d_4)-(a_4+b_4-a_3-b_3)(a_4+c_4-a_2-c_2)},$$

$$(5)\quad p_\infty(B_1) = \frac{(c_2+d_2)(b_4+d_4)+(c_4+d_4)(b_2+d_2)}{(c_2+d_2+c_4+d_4)(b_3+d_3+b_4+d_4)-(a_4+b_4-a_3-b_3)(a_4+c_4-a_2-c_2)}.$$

These equations were derived by solving for $\{u_k\}$ in Equation 1 and then substituting in Equations 2 and 3. It should be noted that these equations are not a general solution to the Markov process described in the first section of this paper, but represent a solution only when the following pair of conditions are satisfied:

$$(6)\quad \begin{aligned} (a_1+a_3)+(b_1+b_3) &= (a_2+a_4)+(b_2+b_4) \\ (a_1+a_2)+(c_1+c_2) &= (a_3+a_4)+(c_3+c_4)\,. \end{aligned}$$

When these conditions do not hold, the solutions for $p_\infty(A_1)$ and $p_\infty(B_1)$ are functions of θ_A and θ_B. Since Equations 4 and 5 are independent of θ_A and θ_B and are strictly functions of the experimental parameter values, it follows that these equations should predict both individual behavior and group mean values. That is, over a sufficiently long series of trials, the observed proportions of responses for an individual subject and the group mean of these quantities both should approach the values predicted by Equations 4 and 5.

Inspection of Table 1 indicates close agreement between observed and predicted values. To check this agreement t tests were run between the observed and predicted values employing the observed standard deviation of the mean as the error term. In all cases the obtained value of t did not approach significance at the .10 level.

A check on the correspondence between individual asymptotic behavior and predicted values is equivalent to evaluating the agreement between observed standard deviations presented in Table 1 and asymptotic variability predicted by the model. Unfortunately, direct computation of the theoretical standard deviation is extremely cumbrous, and we have not obtained an analytical solution. Nevertheless, some results from Monte Carlo runs [3] tentatively suggest that the observed variances are of the proper order to be accounted for in terms of the present model.

Transition Probabilities. Because of the relatively simple mathematical character of stationary Markov processes with a finite number of states, it is possible to ask certain detailed questions. Probably the most immediate question is: How do the aggregate transition matrices for the two experimental groups compare with the theoretical matrix $[p_{ij}]$ that is presented on p. 67?

TABLE 2

Observed Transition Matrices Corresponding to the Theoretical Transition Matrix, Computed over the Last 90 Trials

State	Sure				Mixed			
	$\langle A_1, B_1\rangle$	$\langle A_1, B_2\rangle$	$\langle A_2, B_1\rangle$	$\langle A_2, B_2\rangle$	$\langle A_1, B_1\rangle$	$\langle A_1, B_2\rangle$	$\langle A_2, B_1\rangle$	$\langle A_2, B_2\rangle$
$\langle A_1, B_1\rangle$	.27	.59	.05*	.09*	.22	.63	.05*	.10*
$\langle A_1, B_2\rangle$	.06*	.71	.02*	.21	.05*	.47	.05*	.43
$\langle A_2, B_1\rangle$	.22	.37	.19	.22	.25	.33	.19	.23
$\langle A_2, B_2\rangle$	.14	.43	.07	.36	.12	.39	.15	.34

Table 2 presents values for the observed matrix $[\hat{p}_{ij}]$ computed over the last 90 trials for the two groups. The maximum likelihood estimate $\hat{p}_{ij}$ of the quantity p_{ij} is obtained as follows [1]: (i) Let $N_{ij}(n)$ denote the number of subject pairs in State i on trial $n-1$ and State j on trial n. (ii) Let

$$N_{ij} = \sum_{n=121}^{210} N_{ij}(n);$$

then

$$\hat{p}_{ij} = \frac{N_{ij}}{\sum_{j=1}^{4} N_{ij}}.$$

No statistical test is needed to see that some of the $\hat{p}_{ij}$'s differ significantly from the theoretical values. It suffices to observe that in the theoretical transition matrix, for the set of experimental parameter values employed in both the Mixed and Sure Groups, the last two entries in row 1 and the first and third entries in row 2 are identically zero, whereas in the observed matrices, entries in these cells (denoted by * in Table 2) are in some cases markedly different from zero.

Without regard to a specific model we can ask another highly relevant question about the data: Can the data be more adequately accounted for by a two-stage Markov model which employs information about responses on the previous two trials, as compared with a one-stage model which employs response information about only one preceding trial? For this purpose we use the test described in [1]. The null hypothesis is that $p_{ijk} = p_{jk}$ for $i = 1, 2, 3, 4$, where p_{ijk} is the probability of State k given i and j in succession on the two previous trials, and p_{jk} is the probability of State k simply given State j on the preceding trial. To test this hypothesis the following sum was computed from the aggregate group data:

$$\chi^2 = \sum_{i,j,k} N_{ij}^* \frac{(\hat{p}_{ijk} - \hat{p}_{jk})^2}{\hat{p}_{jk}},$$

where $N_{ij}^* = \sum_k N_{ijk}$, and $\hat{p}_{ijk}$ and N_{ijk} are defined similarly to $\hat{p}_{ij}$ and N_{ij}. If the null hypothesis is true, χ^2 has the usual limiting distribution with $4(4-1)^2 = 36$ degrees of freedom.

The values of χ^2 were 51.9 for the Sure Group and 49.7 for the Mixed Group. In neither case were these values significant at the .05 level. This result indicates that for the present set of data there is no statistically significant improvement in prediction if one knows the response history of the pair of subjects on the previous two trials as against only one preceding trial. The Markov model presented in this paper is formulated as a one-stage process, and this finding supports the model. However, it should be pointed out that this assumption is not necessary for our general theoretical approach.

Game Theory Comparisons. The development of an adequate theory of optimal strategies for non-zero-sum, two-person games has been intensively pursued in the past decade, but as yet no concept of optimality has been proposed which is as solidly based as the minimax concept for zero-sum, two-person games. A natural division of non-zero-sum games is into *cooperative* and *noncooperative* games. In a cooperative game the players are permitted to communicate and bargain before selecting a strategy; in a noncooperative game no such communication and bargaining is permitted. The experimental situation described in this paper corresponds to a noncooperative game.

In certain special non-zero-sum games the highly appealing *sure-thing* principle may be used to select an optimal strategy. In brief, a strategy satisfies the sure-thing principle if no matter what your opponent does, you are at least as well off with this strategy as with any other available to you, and possibly better off. The experimental parameters (a_i, b_i, c_i, d_i) were so selected that for one of the experimental groups, the Sure Group, each subject had available such a strategy, namely A_1 for Player A and B_2 for Player B with probability 1.

Unfortunately, in most non-zero-sum games the sure-thing principle does not lead to a unique optimal strategy, or even to a relatively small class of optimal strategies. In this event, probably the best concept of optimality yet proposed for noncooperative, non-zero-sum games is Nash's notion of an *equilibrium point* ([10], [11]). Roughly speaking, an equilibrium point is a set of strategies, one for each player, such that if all players but one follow their assigned strategies, the remaining player cannot find a better strategy than the one assigned to him. The experimental parameters (a_i, b_i, c_i, d_i) were selected for the second experimental group, the Mixed Group, so that the game had a unique equilibrium point consisting of a mixed strategy for each subject. In particular, Player A should have chosen Response A_1 and Player B Response B_1 with probability 1/5.

Although subjects were not shown the payoff matrix in our experiment, it is a reasonable conjecture that after a large number of trials they would learn enough about the situation to approach an optimal game strategy, i. e., a sure-thing strategy for one group and an equilibrium point for the other. Concerning this conjecture the results for the Sure Group seem conclusive: the optimal strategies of responding A_1 or B_2 with probability 1, for players A and B, respectively, are not even roughly approximated by the observed asymptotic means. Findings for the Mixed Group are also decisive. The results of t tests indicate that the observed asymptotic probability of an A_1 response differs significantly from the equilibrium point strategy of 1/5 beyond the .001 level. And the observed asymptotic probability of a B_1 response differs significantly from the equilibrium point strategy at the .02 level.

Alternative Linear Model. Although the results of this paper have been analyzed in terms of a Markov stimulus-sampling model, it is also possible to use the two-person linear model formulated in Chapter 8, Sec. 9. In particular, the parameters (a_i, b_i, c_i, d_i) used in our experiment satisfy the restrictions of Theorem 9.7 of Chapter 8. The asymptotes given by Equations 9.14 and 9.15 of Chapter 8 are the same as those predicted by the Markov model, which are shown in Table 1. More generally, Equations 4 and 5 of this paper yield the same asymptotes as Equations 9.14 and 9.15 of Chapter 8, respectively; the restriction imposed by Equations 6 on Equations 4 and 5 is the same restriction as the condition $c = g = 0$ of the hypothesis of Theorem 9.7 of Chapter 8.

On the other hand, it is to be emphasized that the prediction of identical asymptotes by the Markov and linear models does not entail the identity of the two models for the experimental situation studied here. The variance of the asymptotic probabilities for individual subjects, as well as any sequential statistic such as the probability of two successive identical responses, is different in the two models. Detailed comparisons are not presented here, because of the difficulty of computing any quantity but the mean response probabilities in the two-person linear model.

Comments

From the standpoint of many social psychologists the experimental situation used in this study is too highly structured in terms of successful performance, and interaction between subjects is too severely restricted. Concepts like those of friendliness, cohesiveness, group pressure, opinion discrepancy, and receptivity, which have been important in numerous recent investigations, play no role in our situation. However, these limitations are offset by some substantial assets. An intrinsically quantitative prediction of behavior in an interaction situation has been derived in a rigorous manner from fundamental principles of reinforcement and association learning. In particular, the only psychological concepts needed for the analysis of our experiment are the classical triad of stimulus, response, and reinforcement. In comparison, studies using common-sense group concepts like those just mentioned have not been quantitative in character, nor have they made any serious headway toward deriving these concepts from any specific psychological theory.

From another viewpoint it is interesting to observe that this study supports results in [3], namely, that various concepts of optimal strategy from the theory of games have not proved useful tools for the prediction of actual behavior. Although this generalization must be qualified by the remark that subjects were not shown the payoff matrix of the game, the relative success of statistical learning models in predicting behavior seems substantial. Still, it is of theoretical interest to find out how much, if at all, explicit knowledge of the payoff matrix on the part of the subjects disturbs the predictive accuracy of the learning model; experimental investigation of this problem is now under way.

REFERENCES

1. Anderson, T. W., and Goodman, L. A. Statistical inference about Markov chains. *Annals of Math. Stat.*, 1957, **28**, 89–110.
2. Atkinson, R. C. A Markov model for discrimination learning. *Psychometrika*, 1958, **23**, 309–22.
3. Atkinson, R. C., and Suppes, P. An analysis of two-person games in terms of statistical learning theory. *J. exp. Psychol.*, 1958, **55**, 369–79.
4. Bush, R. R., and Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.
5. Estes, W. K., and Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–86.
6. Estes, W. K., and Burke, C. J. Application of a statistical model to simple discrimination learning in human subjects. *J. exp. Psychol.*, 1955, **50**, 81–88.
7. Estes, W. K., Burke, C. J., Atkinson, R. C., and Frankmann, J. P. Probabilistic discrimination learning. *J. exp. Psychol.*, 1957, **54**, 233–39.
8. Feller, W. *An introduction to probability theory and its applications.* New York: Wiley, 1950.
9. Fréchet, M. *Recherches théoriques modernes sur le calcul des probabilités*, Vol. II. Paris, 1938.
10. Nash, J. F. The bargaining problem. *Econometrica*, 1950, **18**, 155–62.
11. Nash, J. F. Non-cooperative games. *Annals of Mathematics*, 1951, **54**, 286–95.