AN ANALYSIS OF TWO-PERSON GAME SITUATIONS IN TERMS OF STATISTICAL LEARNING THEORY

BY

RICHARD C. ATKINSON AND PATRICK SUPPES

TECHNICAL REPORT NO. 8

APRIL 25, 1957

PREPARED UNDER CONTRACT Nonr 225(17)

(NR 171-034)

FOR

OFFICE OF NAVAL RESEARCH

BEHAVIORAL SCIENCES DIVISION

APPLIED MATHEMATICS AND STATISTICS LABORATORY

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

AN ANALYSIS OF TWO-PERSON GAME SITUATIONS IN TERMS OF STATISTICAL LEARNING THEORY[1/]

by

RICHARD C. ATKINSON AND PATRICK SUPPES
APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY

This study represents an extension of statistical learning theory to a class of two-person, zero-sum game situations. Because the theory has been mainly developed in connection with experiments dealing with individual learning problems, its predictive success in an experimental area involving interaction between individuals provides an additional measure of the scope of its validity. It should be emphasized that the study reported here was not conceived as providing an empirical test of the adequacy of learning theory as opposed to game theory, for although we use the language of game theory to describe the study the game characteristics of the situation were not apparent to Ss. This point is amplified below.

For the purposes of this experiment a play of a game is a trial. On a given trial each of the two players independently makes a choice between one of two alternatives, that is, he makes one of two possible responses. After the players have indicated their choices, the outcome of the trial is announced to each player.

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

$$\begin{array}{c|cc} & B_1 & B_2 \\ \hline A_1 & a_1 & a_2 \\ A_2 & a_3 & a_4 \end{array}$$

The players are designated A and B. The responses available to player A are A_1 and A_2; similarly, the responses available to player B are B_1 and B_2. If player A selects A_1 and player B selects B_1 then there is a probability a_1 that player A is "correct" and player B is "incorrect," and a probability $1-a_1$ that player A is "incorrect" and player B is "correct." The outcomes of the other three response pairs are identically specified in terms of a_2, a_3 and a_4.

The interaction of the players is limited by two factors: (i) neither player is shown the payoff matrix, (ii) neither player is directly informed of the responses of the other player. Thus, from the standpoint of the general theory of rational behavior (4), S should not regard himself as playing a 2x2 game with known payoff matrix but should view the situation as a multi-stage decision problem against an unknown opponent. However, selection of an optimal strategy in this multi-stage decision problem, is far from a trivial task mathematically, and it is scarcely to be expected that any S would use such a strategy. The virtue of statistical learning theory is that it yields a quantitative prediction of how organisms actually do behave in such situations.

Our theoretical analysis of the behavior of Ss in the situation described is based on two distinct but closely related models. Since a detailed mathematical analysis of these models is presented elsewhere (14), we shall confine ourselves to the most salient facts and omit mathematical proofs.

Linear model.-- The first model is an extension of a linear model developed by Estes and Burke. Experimental tests of this formulation for one person learning situations are reported in (2,9,13). The basic assumption of the model is that response probability on a given trial is a linear function of the probability on the preceding trial. When a response is reinforced its probability increases; the reinforcement of any other response decreases its probability.

In our situation, where two responses are available to each S, we say that if a response occurs and is designated as "correct," then the response is reinforced; if a response occurs and is designated as "incorrect" then the alternative response is reinforced. More specifically, let α_n be the probability of response A_1 on trial n. The rules of change are:

(i) if A_1 is reinforced on trial n then

$$\alpha_{n+1} = (1-\theta_A)\alpha_n + \theta_A$$

(ii) if A_2 is reinforced on trial n then

$$\alpha_{n+1} = (1-\theta_A)\alpha_n$$

where $0 < \theta_A \leq 1$. Identical rules are specified for β_n, the probability of a B_1 response, in terms of θ_B.

The following pair of recursive equations can then be derived for the mean probabilities $\overline{\alpha}_n$ and $\overline{\beta}_n$, where $\overline{\gamma}_n$ is the mean probability of the joint event that on trial n player A will make response A_1 and player B response B_1.

$$\overline{\alpha}_{n+1} = (1-2\theta_A + \theta_A a_2 + \theta_A a_4)\overline{\alpha}_n + \theta_A(a_4-a_3)\overline{\beta}_n + \theta_A(1-a_4)$$
$$+ \theta_A(a_1+a_3 - a_2-a_4)\overline{\gamma}_n$$

$$\bar{\beta}_{n+1} = (1-\theta_B a_3 - \theta_B a_4)\bar{\beta}_n + \theta_B(a_2-a_4)\bar{\alpha}_n + \theta_B a_4$$
$$+ \theta_B(a_3+a_4 - a_1-a_2)\bar{\gamma}_n.$$

It may be shown that $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$, the asymptotic probabilities in the sense of Cesaro (11) 2/, exist and are independent of the initial probabilities α_o, β_o, γ_o. However, in general these asymptotic quantities depend on θ_A and θ_B, and no simple results are obtainable for the quantities individually. On the other hand, an interesting linear relation between $\bar{\alpha}$ and $\bar{\beta}$ which is independent of $\bar{\gamma}$, θ_A and θ_B can be derived, namely:

$$(1)\quad [(a_3+a_4 - a_1-a_2) + (a_1a_2 - a_3a_4)]\bar{\alpha} = (a_1a_3 - a_2a_4)\bar{\beta} + \frac{1}{2}(a_3+a_4 - a_1-a_2)$$
$$+ a_4(a_2-a_3).$$

We have labeled the line determined by this equation the interaction line since the exact point on the line specifying the asymptotic probabilities $\bar{\alpha}$ and $\bar{\beta}$ is a function of both θ_A and θ_B. It is interesting to observe that in the corresponding one-person learning situation the interaction line degenerates to a point, while in the three-person situation we obtain an interaction surface.

Finite Markov model.-- In this model we describe an organism as being in one of two states. If in state 1, he will make response 1; and if in state 2, response 2. Thus in our situation, on any trial n, the two players are described in terms of the following four states: $\langle 1,1 \rangle$, $\langle 1,2 \rangle$, $\langle 2,1 \rangle$, and $\langle 2,2 \rangle$ where the first member of a couple indicates the state of player A and the second, the state of player B. For example $\langle 2,1 \rangle$ means that player A will make response A_2 and player B will make response B_1. We postulate that the change of states from one trial to the next is Markovian, and use the following analysis to derive the transition matrix (10,11) of the process.

When one of players A's responses is reinforced on trial n there is (i) a probability θ_A that the organism is affected by the reinforcing event so that on trial $n+1$ he will make the response reinforced on trial n and (ii) a probability $1-\theta_A$ that the organism is not affected by the reinforcing event and consequently repeats, on trial $n+1$, the response made on trial n. Identical rules describe the process for player B in terms of θ_B.

For this set of assumptions and the payoff probabilities a_1, a_2, a_3 and a_4, the transition matrix describing the learning process can be derived and is as follows:

$$
\begin{array}{c|cccc}
 & \langle 1,1 \rangle & \langle 1,2 \rangle & \langle 2,1 \rangle & \langle 2,2 \rangle \\
\hline
\langle 1,1 \rangle & a_1(\theta_A-\theta_B) + (1-\theta_A) & a_1\theta_B & (1-a_1)\theta_A & 0 \\
\langle 1,2 \rangle & a_2\theta_B & a_2(\theta_A-\theta_B) + (1-\theta_A) & 0 & (1-a_2)\theta_A \\
\langle 2,1 \rangle & (1-a_3)\theta_A & 0 & a_3(\theta_A-\theta_B) + (1-\theta_A) & a_3\theta_B \\
\langle 2,2 \rangle & 0 & (1-a_4)\theta_A & a_4\theta_B & a_4(\theta_A-\theta_B) + (1-\theta_A)
\end{array}
$$

Rows designate the state on trial n and columns the state on trial $n+1$.

Thus $(1-a_3)\theta_A$, the entry in row 3, column 1, is the conditional probability of being in state $< 1,1 >$ on trial $n+1$ given that the pair of Ss was in state $< 2,1 >$ on trial n, for we have:

$$(1-a_3)\theta_A = \theta_A\theta_B(1-a_3) + \theta_A(1-\theta_B)(1-a_3) + (1-\theta_A)\theta_B\cdot 0 + (1-\theta_A)(1-\theta_B)\cdot 0 .$$

From these one stage transition probabilities we obtain an explicit solution for the Cesaro asymptotic probabilities of an A_1 and B_1 response; as in the case of the linear model we denote these quantities $\overline{\alpha}$ and $\overline{\beta}$ respectively. The general equations for $\overline{\alpha}$ and $\overline{\beta}$ are too lengthy to reproduce here but certain results are noteworthy. It can be shown that $\overline{\alpha}$ and $\overline{\beta}$ are related by the identical interaction line determined by Equation (1) of the linear model. For the Markov model, however, we can in addition prove that the point on the interaction line describing a particular pair of Ss' asymptotic behaviors is uniquely determined by the ratio of θ_A to θ_B. Further, even without a knowledge of θ_A and θ_B (i.e., for any combination of θ_A and θ_B) we can specify a fairly narrow interval on the interaction line within which $\overline{\alpha}$ and $\overline{\beta}$ must fall.

Particular cases of the theoretical analysis may be illustrated by examining predictions for the parameter values employed in this experiment. Three sets of a_i values were used corresponding to three classical cases of 2×2 games in the theory of zero-sum, two person games (12).

The first case is labeled the Mixed Group, since both players have mixed minimax strategies. The a_i values are given by the payoff matrix

$$\begin{array}{c|cc} & B_1 & B_2 \\ \hline A_1 & 1/3 & 1 \\ A_2 & 1/2 & 1/6 \end{array}$$

The minimax strategy for player A is to choose A_1 with probability $1/3$, and the minimax strategy for B is to choose B_1 with probability $5/6$. In the Markov model

$$(2) \qquad \bar{\alpha} = .600$$

$$(3) \qquad \bar{\beta} = \frac{35(\theta_A/\theta_B) + 22}{50(\theta_A/\theta_B) + 40} .$$

Note that $\bar{\alpha}$ is independent of θ_A/θ_B. From (3) we obtain as bounds on $\bar{\beta}$:

$$(4) \qquad .550 < \bar{\beta} < .700.$$

If we assume $\theta_A = \theta_B$ then $\bar{\beta} = .633$. For this case the interaction line is the line satisfying (2).

The second case is labelled the <u>Pure</u> Group, since both players have pure minimax strategies. The particular values are given by the matrix

$$\begin{array}{c|cc} & B_1 & B_2 \\ \hline A_1 & 1/2 & 1 \\ A_2 & 1/2 & 1/4 \end{array}$$

Here $a_1 = 1/2$ is a saddle point of the matrix and from the standpoint of game theory the optimal strategy for player A is to play A_1 with probability 1 and for B to play B_1 with probability 1. In the Markov model

$$(5) \qquad \bar{\alpha} = .667$$

$$(6) \qquad \bar{\beta} = \frac{6(\theta_A/\theta_B) + 5}{9(\theta_A/\theta_B) + 9} .$$

As in the previous case, $\bar{\alpha}$ is independent of θ_A/θ_B and the interaction line is the line satisfying (5). From (6) we obtain as bounds on $\bar{\beta}$:

$$(7) \qquad .555 < \bar{\beta} < .667.$$

If we assume that $\theta_A = \theta_B$ then $\bar{\beta} = .611$.

The third case is labelled the Sure Group since both players have sure-thing strategies (i.e., given the payoff matrix one of the two responses available to each player is at least as good or better than the other response regardless of what his opponent does). The parameter values are given by the matrix

	B_1	B_2
A_1	1/2	1
A_2	1/4	3/4

The sure-thing strategies for players A and B are A_1 and B_1 respectively. In the Markov model

$$(8) \qquad \bar{\alpha} = \frac{5(\theta_A/\theta_B) + 15}{7(\theta_A/\theta_B) + 23}$$

$$\bar{\beta} = \frac{5(\theta_A/\theta_B) + 16}{7(\theta_A/\theta_B) + 23}, \tag{9}$$

and as bounds we have:

$$.652 < \bar{\alpha} < .711 \tag{10}$$

$$.696 < \bar{\beta} < .711. \tag{11}$$

If we assume that $\theta_A = \theta_B$ then $\bar{\alpha} = .667$ and $\bar{\beta} = .700$. For this case the interaction line is determined by the equation:

$$3\bar{\alpha} = 10\bar{\beta} - 5. \tag{12}$$

Method

Subjects.-- The <u>Ss</u> were 120 undergraduates obtained from introductory courses in psychology and philosophy at Stanford University. They were randomly assigned to the Mixed, Pure, and Sure Groups with the restriction that there were 20 pairs of <u>Ss</u> in each group.

Apparatus.-- The <u>Ss</u>, run in pairs, sat at opposite ends of an 8ft. by 3ft. table. Mounted vertically on the table top facing each <u>S</u> was a 50in. wide by 30in. high black panel placed 22in. from the end of the table. The <u>E</u> sat between the two panels and was not visible to either <u>S</u>. The apparatus, as viewed from the <u>S</u>'s side, consisted of two silent operating keys mounted 8in. apart on the table top and 12in. from the end of the table; upon the panel, three milk-glass panel lights were mounted. One of these lights, which served as the signal for S to respond, was centered between the keys at a height of 17in. from the table top. Each of the two remaining lights, the reinforcing signals, was at a height of 11in. directly above one of the keys. The presentation and duration of the lights were automatically controlled.

Procedure.-- Ss were read the following instructions: "We always run Ss in pairs because this is the way the equipment has been designed and also because it is the most economical procedure. Actually, however, you are both working on two completely different and independent problems.

"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, and should have pressed the other key. At times you may feel frustrated or irritated if you cannot understand what the experiment is all about. Nevertheless, continue trying to make the very best prediction you can on each trial."

For each pair of Ss, one person was randomly selected as player A and the other player B. Further, for each S 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 three 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.

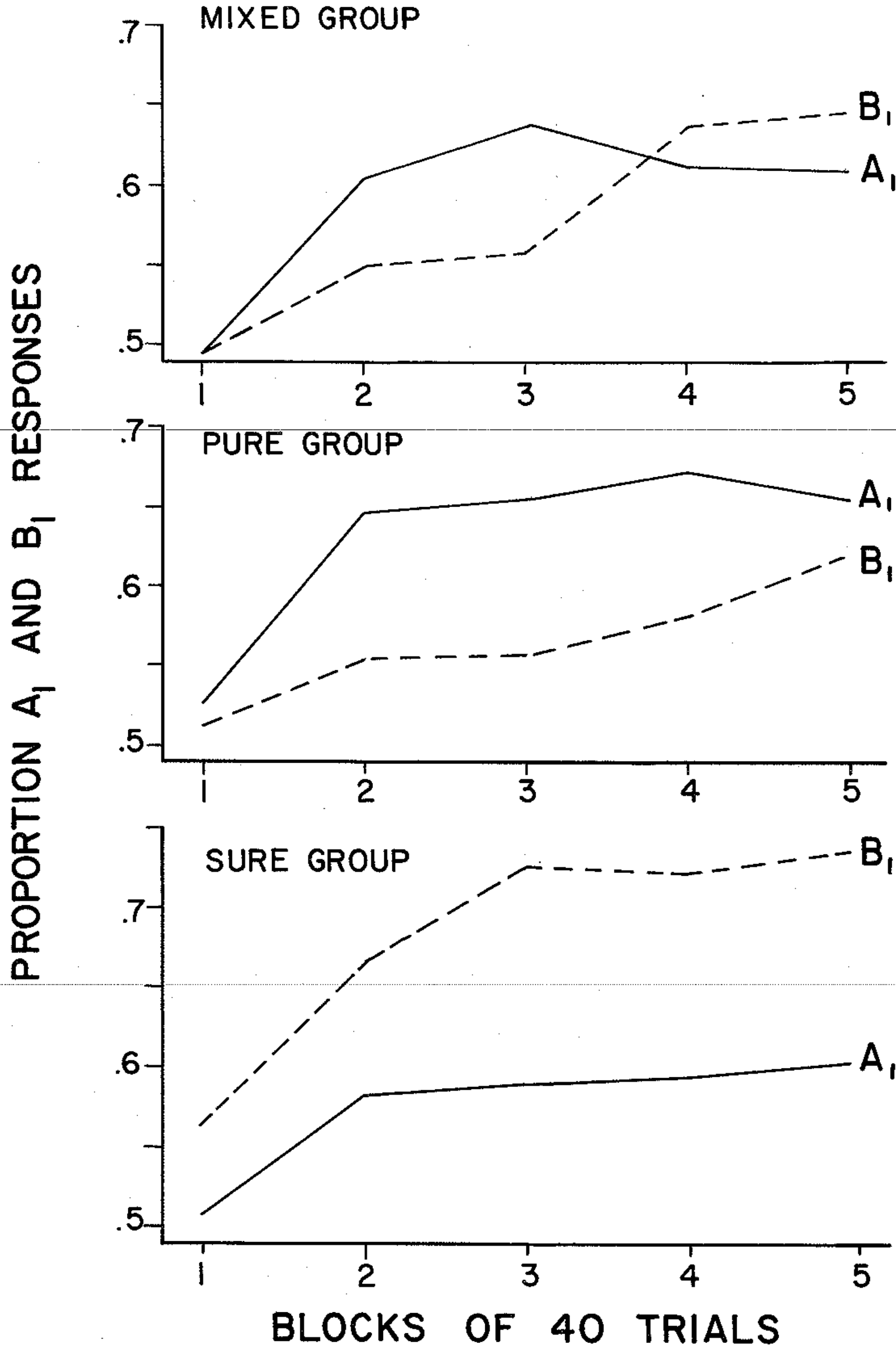
Following the instructions, 200 trials were run in continuous sequence. For each pair of Ss sequences of reinforcing lights were generated in accordance with assigned values of a_i and observed responses.

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

At the end of the session each S was asked to describe what he thought was involved in the experiment. Only one S indicated that he believed the reinforcing events depended in any way on a relationship between his responses and the other players' responses. His record and that of his partner were eliminated from the analysis and replaced by another pair.

Results and Discussion

Mean learning curves and asymptotic results.-- Figure 1 provides a


MIXED GROUP
PURE GROUP
SURE GROUP
A_1
B_1
.7
.6
.5
1
2
3
4
5
PROPORTION A_1 AND B_1 RESPONSES
BLOCKS OF 40 TRIALS

description of behavior over all trials of the experiment. In this figure the mean proportions of A_1 and B_1 responses in succesive blocks of 40 trials are given for the sequence of 200 trials. An inspection of this figure indicates that responses are fairly stable over the last 100 trials except possibly for B_1 responses in the Pure Group. To check the stability of response behavior for individual data, $\underline{t}$ s for paired measures were computed between response proportions for the first and last halves of the final block of 60 trials. In all cases the obtained values of $\underline{t}$ fall short of significance at the .05 level.

It appears reasonable to assume that a constant level of responding has been reached; consequently the proportions computed over the last 60 trials were used as an estimate of $\overline{\alpha}$ and $\overline{\beta}$. Table 1 presents the observed mean proportions of A_1 and B_1 responses in the last 60 trial block and the standard deviations associated with these means. Each entry is based on $N = 20$. The values predicted by the Markov model for $\theta_A = \theta_B$ are also presented.

Inspection of Table 1 indicates that predicted and observed results are extremely close for the Mixed and Pure Groups; $\underline{t}$ tests of the difference between these values do not approach significance at the .05 level. For the Sure Group the difference between player B's observed and theoretical values is also not significant; but for player A, the difference is significant. Specifically, the observed proportion of A_1 responses for the Sure Group is less than predicted. Note, however, that we may relinquish the assumption that $\theta_A = \theta_B$ and, given the boundary conditions specified by Equations (10) and (11), determine for

Table 1. Predicted and Observed Mean Proportions of A_1 and B_1 Responses over the Last Block of 60 Trials

	A_1			B_1		
	Predicted	Observed	s	Predicted	Observed	s
Mixed	.600	.605	.0794	.633	.649	.0874
Pure	.667	.670	.0832	.611	.602	.0634
Sure	.667	.606	.1005	.700	.731	.0760

the Sure Group the point on the interaction line (see Equation 12) which is nearest the observed point. This nearest point is $\bar{\alpha} = .652$ and $\bar{\beta} = .696$. For this point the difference between observed and theoretical values is not significant at the .05 level for either A_1 or B_1 responses.

Game theory comparisons.-- It is of interest to compare observed values with the game-theoretic optimal strategies discussed earlier, for it can be reasonably maintained that even though Ss do not know the pay-off matrix, after a large number of trials they have learned enough about the situation to approach an optimal game strategy. Concerning such a conjecture the results for the Pure and Sure Groups seem decisive: the optimal game strategies of responding A_1 or B_1 with probability 1, for player A or B respectively, is not even crudely approximated by the observed means. Moreover, the maximum individual value in each group of 20 Ss does not approach 1; for the Pure Group max α is .80 and max β is .71, while for the Sure Group max α is .77 and max β is .84.

The results for the Mixed Group also fail to support the hypothesis that Ss, in the long run, will approach an optimal game strategy. The observed $\bar{\alpha}$ of .605 and $\bar{\beta}$ of .649 both differ significantly from their respective minimax strategies of 1/3 and 5/6 at beyond the .001 level.

Several questions are suggested by these comparisons with game theory that are pertinent to a theory of small groups. First, would the learning theory predictions be less applicable and the optimal game strategies more closely approximated if Ss are explicitly told that they are competing with each other? Subsequent experimental work (3) indicates that the answer to this question is probably negative. Second, would optimal game strategies be more closely approximated if Ss were run for a very large number of trials over a period of several days? What evidence there is on this question from individual learning situations (7, 8, 9) tends to support the hypothesis that the long run mean probabilities would stay close to the learning theory predictions. However, detailed experimental investigation would be worthwhile. Third, would the present experimental results be affected if Ss were paid for correct responses and penalized monetarily for incorrect responses? The models formulated in the first part of this paper are not rich enough in conceptual content to express formally possible effects of different types of reinforcing events. Fourth, will the obvious generalization of the two models to the interaction of more than two Ss be experimentally substantiated, and how will observed response probabilities compare with various proposed "solutions" of n-person games?

Adequacy of Markov model.-- 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 about our data from the standpoint of Markov models. Probably the most direct question is: how do the aggregate transition matrices for each of the three experimental groups compare with the theoretical transition matrix derived in the first part of the paper? Table 2 presents the observed matrices computed over the last 100 trials. Since each group contained 20 pairs of Ss each matrix is based on 2000 observations. No statistical test is needed to see that the observed matrices differ significantly from the theoretical matrix. It is sufficient to observe that in the theoretical matrix (for all sets of

Table 2. Observed Transition Matrices Corresponding to the Theoretical Transition Matrix Specified by the Markov Model. Computed over the Last 100 Trials.

	Mixed				Pure				Sure			
	< 1,1 >	< 1,2 >	< 2,1 >	< 2,2 >	< 1,1 >	< 1,2 >	< 2,1 >	< 2,2 >	< 1,1 >	< 1,2 >	< 2,1 >	< 2,2 >
< 1,1 >	.37	.22	.30	.11	.38	.27	.24	.11	.43	.18	.29	.10
< 1,2 >	.54	.25	.15	.06	.50	.31	.11	.08	.52	.19	.22	.07
< 2,1 >	.35	.16	.30	.19	.30	.20	.29	.21	.47	.12	.31	.10
< 2,2 >	.28	.34	.17	.21	.31	.36	.16	.17	.27	.17	.37	.19

parameter values a_i) the antidiagonal is identically zero, but in the observed matrices every entry on the antidiagonals is markedly different from zero. As a matter of fact it would be surprising to find a very close agreement between the theoretical and observed matrices, for the theoretical matrix was derived from exceedingly simple assumptions. From a psychological standpoint our Markov model can be interpreted, for a given player, as a one-element stimulus model, where the stimulus is sampled with probability 1 on every trial and conditioned to the reinforced response class with probability θ. It seems unlikely that the detailed pattern of Ss responses could be accounted for by a single stimulus element.

For experimental situations involving more than one organism even the extension to a two-element stimulus model is not trivial from the standpoint of computing the simplest quantities desired, namely, asymptotic response probabilities. For example, if in the one-element model we identified the stimulus element as the signal light, one natural two-element model is to identify two successive signals as the two stimuli. The Markov process derived from this assumption has, for our experiment, sixteen states.

Fortunately, without examining a specific two-stage Markov model we can ask one highly relevant question about our data: can the data be more adequately accounted for by a two-stage model which employs information about the organism on the previous two trials as compared with a one-stage model which employs information about only one preceding trial? For this purpose we use the χ^2-test described in (1). The null hypothesis is that $p_{ijk} = p_{jk}$ for $i=1,\ldots,4$ where p_{ijk} is the probability of state k given successively states i and j 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 for aggregate group data:

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

where $n^*_{ij} = \sum_k n_{ijk}$. If the null hypothesis is true, χ^2 has the usual limiting distribution with $4(4-1)^2 = 36$ degrees of freedom.

The obtained values of χ^2 were 81.8 for the Mixed Group, 50.5 for the Pure Group and 52.9 for the Sure Group. For the Pure and Sure Groups the value of χ^2 is not significant at the .05 level, for the Mixed Group it is. Independent of any specific model these results indicate that for two of the three groups the learning process is fairly well approximated by a one-stage Markov process. Moreover, it is to be noted that the significant χ^2 for the aggregated data of the Mixed Group does not entail that individual pairs of Ss were not one-stage Markovian in their responses, for the sum (in the sense pertinent here) of several Markov processes is not necessarily a Markov process. The small number of observations for a given pair of Ss ruled out a separate χ^2-test for each pair.

If, on the other hand, we accept the approximate one-stage Markovian character of the learning process studied in this experiment, and ask if this process is stationary in the sense that the observed transition probabilities are constant over all trials, the answer is decisively negative. In a χ^2 analysis (1) of the aggregate observed transition matrix for the first 100 trials compared with the last 100 trials, the difference was significant at the .01 level for all three groups. These results suggest that non-stationary, single element models need to be explored in addition to an analysis of stationary multi-stimulus element models.

Observed and predicted variances in the linear model.-- The close agreement between predicted and observed mean asymptotic responses suggests a check of the linear model against another measure of behavior. Specifically, we were interested in checking the variance predicted by the model against the experimental results on variability presented in Table 1. The observed standard deviations in this table relate to the proportions of A_1 and B_1 responses in blocks of 60 trials, comparable theoretical quantities will be designated $\sigma(A_1)$ and $\sigma(B_1)$ respectively.

Unfortunately direct analytical computation of $\sigma(A_1)$ and $\sigma(B_1)$ seems impossible. Consequently it was necessary to resort to "Monte Carlo methods" (5). The basic idea of the approach is to construct a system which follows the rules specified by the theory and then make observations on the behavior of the system. By taking a large number of such observations one obtains precise estimates of theoretical quantities. In our case, what might be considered a hypothetical organism was built by programming an I.B.M. type 650 digital computor so that its sequence of commands corresponded to the operations specified by the linear model.

Employing this procedure, estimates of $\sigma(A_1)$ and $\sigma(B_1)$ were obtained for various values of θ_A and θ_B. Results from other experiments (2, 9, 13, 15) suggested that θ values for the present study would undoubtedly be bounded between .01 and .50. Hence combinations of .01, .10, and .50 were used in the computation; a finer gradation of values would have been desirable but the cost of computor time made this prohibitive. The results of the Monte Carlo runs are presented in Table 3.

A comparison of Tables 1 and 3 indicates that, for all cases, the observed variability is greater than predicted by the model. Even the most favorable comparisons between observed and predicted values prove to be significantly different at the .05 level when a χ^2-test of variances is employed. The finding that the linear model tends to underestimate observed variability is not surprising in view of similar results from other experiments employing linear operator models to account for individual learning data.

Table 3. Monte Carlo Estimates of $\sigma(A_1)$ and $\sigma(B_1)$ for Various Values of θ_A and θ_B.

θ_A	θ_B	Mixed		Pure		Sure	
		$\sigma(A_1)$	$\sigma(B_1)$	$\sigma(A_1)$	$\sigma(B_1)$	$\sigma(A_1)$	$\sigma(B_1)$
.01	.01	.030	.031	.033	.024	.024	.021
.01	.10	.021	.042	.034	.051	.021	.047
.01	.50	.027	.061	.026	.053	.030	.046
.10	.01	.049	.031	.058	.022	.051	.021
.10	.10	.050	.052	.059	.036	.050	.035
.10	.50	.060	.057	.060	.043	.065	.053
.50	.01	.055	.024	.076	.018	.071	.021
.50	.10	.062	.037	.064	.051	.049	.045
.50	.50	.066	.045	.071	.046	.074	.036

Summary

The study deals with an analysis of a zero-sum, two-person game situation in terms of statistical learning theory and game theory.

Ss were run in pairs for 200 trials. A single play of the game is treated as a trial. On a trial each player makes a choice between one of two alternative responses; after the players have made their response, the outcome of the trial is announced. The responses available to player A are designated A_1 and A_2; similarly the responses available to player B are B_1 and B_2. If player A selects A_1 and player B selects B_1, then there is a probability a_1 that player A is "correct" and player B is "incorrect" and a probability $1-a_1$ that player B is "correct" and player A is "incorrect". The outcome of the other three response pairs is identically specified in terms of a_2, a_3, and a_4. Ss were instructed to maximize the number of correct responses.

Three groups were run, each employing a different set of a_i values. The selection of these values was determined by game-theoretic considerations; that is, a group had available either a sure-thing strategy, a pure minimax strategy or a mixed minimax strategy.

Analysis of the data was in terms of two different but related stochastic models for learning and game theory. Specifically the following detailed comparisons of data and theory were made: (a) mean asymptotic response probabilities, (b) one and two stage transition probabilities, and (c) variances associated with asymptotic response probabilities.

Footnotes

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. The authors are indebted to W. K. Estes for several stimulating discussions of the ideas on which this experiment is based.

2/ To be explicit, for any k

$$\overline{\alpha} = \lim_{n \to \infty} \frac{1}{n} \sum_{i=k}^{n} \overline{\alpha}_i ,$$

and similarly, for $\overline{\beta}$ and $\overline{\gamma}$.

References

1. Anderson, T. W., & Goodman, L. A. Statistical inference about Markov chains. unpublished manuscript, Statistical Research Center, University of Chicago, 1956.

2. Atkinson, R. C. An analysis of the effect of non-reinforced trials in terms of statistical learning theory. _J. exp. Psychol._, 1956, 52, 28-32.

3. Atkinson, R. C., & Suppes, P. An analysis of non-zero sum games in terms of a Markov model for learning. Technical Report No. 9, Contract NR 171-034, Stanford University.

4. Blackwell, D., & Girshick, M. A. _Theory of games and statistical decisions_. New York: Wiley, 1954.

5. Bush, R. R., & Mosteller, F. _Stochastic models for learning_. New York: Wiley, 1955.

6. Estes, W. K., & Burke, C. J. A theory of stimulus variability in learning. _Psychol. Rev._, 1953, 60, 276-286.

7. Estes, W. K., & Burke, C. J. Application of a statistical model to simple discrimination learning in human subjects. _J. exp. Psychol._, 1955, 50, 81-88.

8. Estes, W. K., Burke, C. J., Atkinson, R. C., & Frankmann, J. P. Probabilistic discrimination learning. _J. exp. Psychol._, 1957, in press.

9. Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. _J. exp. Psychol._, 1954, 47, 225-234.

10. Feller, W. _An introduction to probability and its applications_. New York: Wiley, 1950.

Reference (cont.)

11. Frechet, M. Recherches theoriques modernes sur le calcul des probabilites, vol. 2, Paris, 1938.

12. McKinsey, J. C. C. Introduction to the theory of games. New York: McGraw-Hill, 1952.

13. Neimark, Edith D. Effect of type of non-reinforcement and number of alternative responses in two verbal conditioning situations. J. exp. Psychol., 1956, 52, 209-220.

14. Suppes, P. & Atkinson, R. C. A Markovian learning model for two-person games. Technical Report No. 10, Office of Naval Research Contract NR 171-034, Stanford University.

15. Toda, M. Guessing sequence under various conditions of payoff. Jap. Psychol. Research, 1956, 4, 11-22.

STANFORD UNIVERSITY
Technical Reports Distribution List
Contract Nonr 225(17)
(NR 171-034)

ASTIA Documents Service Center Knott Building Dayton 2, Ohio	5
Commanding Officer Office of Naval Research Branch Office Navy #100, Fleet Post Office New York, New York	35
Director, Naval Research Laboratory Attn: Technical Information Officer Washington 25, D. C.	6
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, New York	1
Office of Naval Research Branch Office 1000 Geary Street San Francisco 9, Calif.	1
Office of Naval Research Branch Office 1030 Green Street Pasadena 1, Calif.	1
Office of Naval Research Branch Office Tenth Floor The John Crerar Library Building 86 East Randolph Street Chicago 1, Illinois	1
Office of Technical Services Department of Commerce Washington 25, D. C.	1
Office of Naval Research Mathematics Division, Code 430 Department of the Navy Washington 25, D. C.	1
Office of Naval Research Logistics Branch, Code 436 Department of the Navy Washington 25, D. C.	1
Operations Research Office 7100 Connecticut Avenue Chevy Chase, Maryland Attn: The Library	1
The Rand Corporation 1700 Main Street Santa Monica, Calif. Attn: Dr. John Kennedy	1
The Logistics Research Project The George Washington Univ. 707 - 22nd Street, N. W. Washington 7, D. C.	1
Dr. R. F. Bales Department of Social Relations Harvard University Cambridge, Massachusetts	1
Dr. Alex Bavelas Massachusetts Institute of Technology Cambridge, Massachusetts	1
Dr. Donald Campbell Department of Psychology Northwestern University Evanston, Illinois	1
Dr. Clyde H. Coombs Bureau of Psychological Services University of Michigan 1027 E. Huron Street Ann Arbor, Michigan	1

Dr. Mort Deutsch
Graduate School of Arts
& Sciences
New York University
Washington Square
New York 3, New York 1

Dr. Francis J. DiVesta
Department of Psychology
Syracuse University
123 College Place
Syracuse, New York 1

Dr. Leon Festinger
Department of Psychology
Stanford University 1

Dr. Murray Gerstenhaber
University of Pennsylvania
Philadelphia, Pennsylvania 1

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

Dr. Harry Helson
Department of Psychology
University of Texas
Austin, Texas 1

Dr. William E. Kappauf
Department of Psychology
University of Illinois
Urbana, Illinois 1

Dr. Leo Katz
Department of Mathematics
Michigan State College
East Lansing, Michigan 1

Dr. Duncan Luce
Bureau of Applied Social Research
Columbia University
New York 27, New York 1

Dr. Nathan Maccoby
Boston University Graduate School
Boston 15, Massachusetts 1

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

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

Dr. Helen Peak
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. George Saslow
Department of Neuropsychiatry
Washington University
640 South Kingshighway
St. Louis, Missouri 1

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

Dr. Marvin Shaw
The Johns Hopkins University
Mergenthaler Hall
Baltimore, Maryland 1

Dr. Herbert Solomon
Teachers College
Columbia University
New York, New York 1

Dr. F. F. Stephan
Box 337
Princeton University
Princeton, New Jersey 1

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

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

Dr. E. Paul Torrance
Survival Research Field Unit
Crew Research Laboratory
AFP & TRC
Stead Air Force Base
Reno, Nevada 1

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

Professor K. J. Arrow
Department of Economics
Stanford University
Stanford, Calif. 1

Professor M. Flood
Willow Run Laboratories
Ypsilanti, Mich. 1

Professor Jacob Marschak
Box 2125, Yale Station
New Haven, Conn. 1

Professor Oskar Morgenstern
Department of Economics &
Social Institutions
Princeton University
Princeton, New Jersey 1

Professor Nicholos Rashevsky
University of Chicago
Chicago 37, Illinois 1

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

Professor Tsunehiko Watanabe
Economics Department
Stanford University
Stanford, California 1

Professor Alan J. Rowe
Management Sciences Research
Project
University of California
Los Angeles 24, California 1

Professor L. J. Savage
Committee on Statistics
University of Chicago
Chicago, Illinois 1

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

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

Professor A. W. Tucker
Department of Mathematics
Princeton University, Fine Hall
Princeton, New Jersey 1

Professor J. Wolfowitz
Department of Mathematics
Cornell University
Ithaca, New York 1

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

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

Professor R. B. Braithwaite
King's College
Cambridge, England 1

Professor Maurice Fréchet
Institut H. Poincaré
11 Rue P. Curie
Paris 5, France 1