Journal of Experimental Psychology
Vol. 55, No. 4, 1958

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

RICHARD C. ATKINSON[2] 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; although we use the language of game theory to describe the study, the game characteristics of the situation were not apparent to *S*s. 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 pay-off matrix.

	B_1	B_2
A_1	a_1	a_2
A_2	a_3	a_4

The players are designated A and B. The responses available to A are A_1 and A_2; similarly, the responses available to B are B_1 and B_2. If A selects A_1 and B selects B_1 then there is a probability a_1 that A is "correct" and B is "incorrect," and a probability $1 - a_1$ that A is "incorrect" and B is "correct." These two joint events are exhaustive since it is required that exactly one player is correct on each trial. 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: (*a*) neither player is shown the pay-off matrix, (*b*) 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 2×2 game with known pay-off 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 *S*s in the situation described is based on two distinct but closely related models. Since a detailed mathematical analysis of these models will be presented elsewhere, the present statement will concern only 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 (**6**). Experimental tests of this formulation for one-person learning situations have been reported

[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] Now at the University of California, Los Angeles.

(**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 the present situation, where two responses are available to each *S*, 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: (*a*) if A_1 is reinforced on Trial *n*, then $\alpha_{n+1} = (1 - \theta_A)\alpha_n + \theta_A$; (*b*) 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 $\bar{\alpha}_n$ and $\bar{\beta}_n$, where $\bar{\gamma}_n$ 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.

$$\begin{aligned}\bar{\alpha}_{n+1} = {} & (1 - 2\theta_A + \theta_A a_2 + \theta_A a_4)\bar{\alpha}_n \\ & + \theta_A(a_4 - a_3)\bar{\beta}_n + \theta_A(1 - a_4) \\ & + \theta_A(a_1 + a_3 - a_2 - a_4)\bar{\gamma}_n\end{aligned}$$

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

It may be shown that $\bar{\alpha}$, $\bar{\beta}$ and $\bar{\gamma}$, the asymptotic probabilities in the sense of Cesaro (**11**),[3] exist and are independent of the initial probabilities α_0, β_0, γ_0. 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:

$$\begin{aligned}[(a_3 + a_4 - a_1 - a_2) + (a_1 a_2 - a_3 a_4)]\bar{\alpha} \\ = (a_1 a_3 - a_2 a_4)\bar{\beta} + \tfrac{1}{2}(a_3 + a_4 - a_1 - a_2) \\ + a_4(a_2 - a_3). \quad (1)\end{aligned}$$

[3] To be explicit,

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

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

The line determined by this equation has been labeled 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 an interaction surface is obtained.

Finite Markov model.—In this model the simplifying assumption is made that on all trials a player's response behavior is determined by a single stimulus—that is, the event associated with the onset of a trial. The *S* is described as being in one of two possible states: (*a*) if in State 1, the stimulus is conditioned to Response 1 and, in the presence of the stimulus, Response 1 will be elicited; (*b*) if 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 following four states: $<1,1>$, $<1,2>$, $<2,1>$ and $<2,2>$ where the first member of a couple indicates the state of Player A and the second, the state of Player B. For example, $<2,1>$ means that Player A will make response A_2 and Player B will make response B_1. It is postulated that the change of states from one trial to the next is Markovian, and the following analysis is used to derive the transition matrix (**10, 11**) of the process.

When one of Player A's responses is reinforced on Trial *n* there is (*a*) a probability θ_A that the stimulus governing Player A's response will be conditioned to the reinforced response and therefore, on Trial $n + 1$ Player A will make the response reinforced on Trial *n* and, (*b*) a probability $1 - \theta_A$ that the conditioned status of the stimulus will remain unchanged and therefore, on Trial $n + 1$ Player A will repeat the response made on Trial *n*. Identical rules describe the process for Player B in terms of θ_B.[4]

[4] The Markov process derived from these assumptions differs in certain respects from that which can be derived from the Estes and Burke stimulus sampling model (6). In their formulation the stimulus is conceptualized as being

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

	<1,1>	<1,2>	<2,1>	<2,2>
<1,1>	$a_1(\theta_A - \theta_B) + (1 - \theta_A)$	$a_1\theta_B$	$(1 - a_1)\theta_A$	0
<1,2>	$a_2\theta_B$	$a_2(\theta_A - \theta_B) + (1 - \theta_A)$	0	$(1 - a_2)\theta_A$
<2,1>	$(1 - a_3)\theta_A$	0	$a_3(\theta_A - \theta_B) + (1 - \theta_A)$	$a_3\theta_B$
<2,2>	0	$(1 - a_4)\theta_A$	$a_4\theta_B$	$a_4(\theta_A - \theta_B) + (1 - \theta_A)$

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 *S*s was in State <2,1> on Trial n, because:

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

From these one-stage transition probabilities an explicit solution is obtained for the Cesaro asymptotic probabilities of an A_1 and B_1 response; as in the case of the linear model these quantities are denoted as $\bar{\alpha}$ and $\bar{\beta}$, respectively. The general equations for $\bar{\alpha}$ and $\bar{\beta}$ are too lengthy to reproduce here but certain results are noteworthy. It can be shown that $\bar{\alpha}$ and $\bar{\beta}$ are related by the identical interaction line determined by Equation 1 of the linear model. For the Markov model, however, it can in addition be proved that the point on the interaction line describing a particular pair of *S*s' asymptotic behaviors is uniquely determined by the ratio of θ_A to θ_B. Further, even without a knowledge of the specific values of θ_A and θ_B one can specify a fairly narrow interval on the interaction line within which $\bar{\alpha}$ and $\bar{\beta}$ must fall by taking the limits of $\bar{\alpha}$ and $\bar{\beta}$ as the ratio θ_A/θ_B approaches zero or becomes large.

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 pay-off matrix

	B_1	B_2
A_1	$\frac{1}{3}$	1
A_2	$\frac{1}{2}$	$\frac{1}{6}$

The minimax strategy for Player A is to choose A_1 with probability $\frac{1}{3}$, and the minimax strategy for B is to choose B_1 with probability $\frac{5}{6}$. In the Markov model

$$\bar{\alpha} = .600 \tag{2}$$

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

Note that $\bar{\alpha}$ is independent of θ_A/θ_B. From Equation 3 one obtains as bounds on $\bar{\beta}$:

$$.550 < \bar{\beta} < .700. \tag{4}$$

If one assumes $\theta_A = \theta_B$, then $\bar{\beta} = .633$. For this case the interaction line is the line satisfying Equation 2.

The second case is labeled the ***Pure Group***, since both players have pure minimax strategies. The particular values

composed of a *large* number of stimulus elements each of which is sampled with probability θ and, once sampled, conditioned to the reinforced response with probability 1. Further, the probability of a response is defined as the proportion of stimulus elements in the sample conditioned to the response. In the model used in this paper it is assumed that the single stimulus is sampled on each trial with probability 1.

are given by the matrix

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

Here $a_1 = \frac{1}{2}$ 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

$$= .667 \tag{5}$$

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

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

$$.555 < \bar{\beta} < .667. \tag{7}$$

If one assumes that $\theta_A = \theta_B$, then $\bar{\beta} = .611$.

The third case is labeled the *Sure Group* since both players have sure-thing strategies (i.e., given the pay-off 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

$$\begin{array}{c|cc} & B_1 & B_2 \\ \hline A_1 & \frac{1}{2} & 1 \\ A_2 & \frac{1}{4} & \frac{3}{4} \end{array}$$

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

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

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

and as bounds one has:

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

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

If one assumes 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 Ss 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 Ss in each group.

Apparatus.—The Ss, run in pairs, sat at opposite ends of an 8 × 3-ft. table. Mounted vertically on the table top facing each S was a 50-in. wide by 30-in. high black panel placed 22 in. from the end of the table. The *E* sat between the two panels and was not visible to either *S*. The apparatus, as viewed from *S*'s side, consisted of two silent operating keys mounted 8 in. apart on the table top and 12 in. 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 17 in. from the table top. Each of the two remaining lights, the reinforcing signals, was at a height of 11 in. directly above one of the keys. The presentation and duration of the lights were automatically controlled. The *Ss* were not visible to one another and could not see each other's responses or panel lights.

Procedure.—The *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 2 sec. 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 was randomly selected as Player A and the other as 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 combination, 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 player's 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 description of behavior over all trials of the experiment. In this figure the mean proportions of A_1 and B_1 responses in successive 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, *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 *t* fall short of significance at the .05 level.

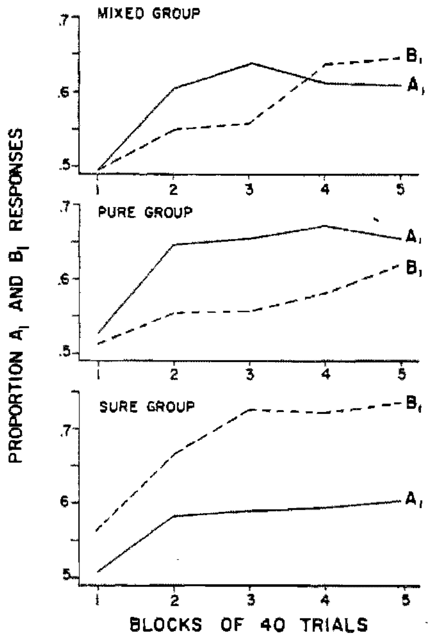

Fig. 1. Observed proportions of A_1 and B_1 responses in blocks of 40 trials for the three experimental groups.

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 $\bar{\alpha}$ and $\bar{\beta}$. Table 1 presents the observed mean proportions of A_1 and B_1 responses in the last 60 trial block and the *SD*'s associated with these means.

TABLE 1

Predicted and Observed Mean Proportions of A_1 and B_1 Responses Over the Last Block of 60 Trials

	Response A_1			Response B_1		
	Pred.	Obs.	*SD*	Pred.	Obs.	*SD*
Mixed	.600	.605	.0794	.633	.649	.0874
Pure	.667	.670	.0832	.611	.602	.0634
Sure	.667	.606	.1005	.700	.731	.0760

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; 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 one may relinquish the assumption that $\theta_A = \theta_B$ and, given the boundary conditions specified by Equations 10 and 11, determine for 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, are 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 $\frac{1}{3}$ and $\frac{2}{3}$ 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 worth while. 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

TABLE 2

Observed Transition Matrices Corresponding to the Theoretical Transition Matrix Specified by the Markov Model

(Computed over the last 100 Trials)

	Mixed Group				Pure Group				Sure Group			
	<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

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 *S*s, 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 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. In particular, it seems unlikely that the detailed pattern of *S*s' responses could be accounted for by a model which conceptualizes the effective stimulus as a single element identical from trial to trial.

For experimental situations involving more than one *S* 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 one identifies 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 the present experiment, 16 states.

Fortunately, without examining a specific two-stage Markov model one can ask one highly relevant question about the present data: Can the data be more adequately accounted for by a two-stage model which employs information about the *S* on the previous two trials as compared with a one-stage model which employs information about only one preceding trial? For this purpose the χ^2 test described by Anderson and Goodman (1) was used. The null hypothesis is that $p_{ijk} = p_{jk}$ for $i = 1, \cdots, 4$ 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 $4(4-1)^2 = 36\,df$.

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 *S*s 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 relatively small number of observations for a given pair of *S*s ruled out a separate χ^2 test for each pair.

If, on the other hand, one accepts the approximate one-stage Markovian character of the learning process studied in this experiment, and asks if this process is stationary in the sense that the observed transition probabilities are constant over all trials, the answer is 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 nonstationary, 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 *SD*'s 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 the present case, what might be considered a hypothetical *S* 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, 14**) 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.

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 Group		Pure Group		Sure Group	
		$\sigma(A_1)$	$\sigma(B_1)$	$\sigma(A_1)$	$\sigma(B_1)$	$\sigma(A_1)$	$\sigma(B_1)$
.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

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.

SUMMARY

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

The *S*s 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. The *S*s 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.

REFERENCES

1. ANDERSON, T. W., & GOODMAN, L. A. Statistical inference about Markov chains. *Ann. math. Statist.*, 1957, **28**, 89–110.
2. ATKINSON, R. C. An analysis of the effect of nonreinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, **52**, 28–32.
3. ATKINSON, R. C., & SUPPES, P. An analysis of a two-person interaction situation in terms of a Markov process. Tech. Rep. No. 9, Contract NR 171-034, Applied Mathematics & Statistics Laboratory, Stanford Univer., 1957.
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, **54**, 233–239.
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 theory and its applications*. New York: Wiley, 1950.
11. Frechet, M. *Recherches théoriques modernes sur le calcul des probabilités*, Vol. 2, Paris: Gauthier-Villars, 1938.
12. McKinsey, J. C. C. *Introduction to the theory of games*. New York: McGraw-Hill, 1952.
13. Neimark, E. D. Effect of type of nonreinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 209–220.
14. Toda, M. Guessing sequence under various conditions of payoff. *Jap. psychol. Res.*, 1956, **4**, 11–22.

(Received April 29, 1957)