Title

Choice Behavior and Reward Structure

Authors

Myers, Jerome L.

Atkinson, Richard C.

Publication Date

1963-05-24

Peer reviewed

CHOICE BEHAVIOR AND REWARD STRUCTURE

by

Jerome L. Myers and Richard C. Atkinson

TECHNICAL REPORT NO. 54

May 24, 1963

PSYCHOLOGY SERIES

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

Abstract

A model for choice behavior under payoff is presented. Predictions of choice probabilities are evaluated for several experiments involving different event probabilities and payoff levels, two and three choices, and contingent and noncontingent reinforcement. An extension of the model to the prediction of response time is also considered.

CHOICE BEHAVIOR AND REWARD STRUCTURE[1]

Jerome L. Myers
University of Massachusetts

Richard C. Atkinson
Stanford University

A growing body of data, from both animal and human experimentation, reflects the importance of magnitude of reinforcement in choice behavior. In this paper, we attempt a quantitative description of the role of this variable. The model under consideration was originally proposed by Atkinson (1962), who showed that it was applicable to both contingent and noncontingent reinforcement, and to any number of response alternatives. We have extended this work by deriving predictions of response times, and of conditional statistics that were not previously presented. We have also considered ways of formulating the model to yield predictions for more general reward-punishment combinations than those previously considered. In addition, the observed and predicted values of a number of measures obtained from several different experiments are displayed in this paper. We hope that the data presented will provide an impetus to the development of alternative models, and that the description of the data by our model will provide a criterion against which to judge other models.

The model that we will consider assumes a population of stimulus elements, each of which is conditioned to one and only one response. It is further assumed that a single element is randomly sampled from the stimulus population on each trial, and that the subject's response depends upon the state of conditioning of the sampled element. These assumptions are common to other stimulus sampling models (Atkinson and Estes, 1963); however the present model differs from its predecessors for it is assumed that an element may be at one of two stages of conditioning, either weakly or strongly conditioned to a response. This modification of stimulus sampling theory provides the basis for a fairly general analysis of reinforcement variables.

Axioms

Consider an experiment in which on each trial the subject must select one of r mutually exclusive and exhaustive responses $(A_1,\ldots,A_i,\ldots,A_r)$. The i^{th} response corresponds to the prediction of the i^{th} member of a set of r mutually exclusive and exhaustive events $(E_1,\ldots,E_i,\ldots,E_r)$. Associated with each response-event combination is some outcome such as the gain or loss of an amount of money. In the experiments that we examine the $r \times r$ set of outcomes is constant over trials. We will first consider axioms for the special case in which each $A_i - E_i$ combination is followed by the same gain, and all other response-event combinations are followed by the same loss. This situation may be represented by the following payoff matrix

$$
\begin{array}{c}
 \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_r
\end{array}
\begin{array}{c}
\begin{array}{ccccc} E_1 & \cdots & E_i & \cdots & E_r \end{array} \\
\begin{bmatrix}
w & \cdots & -x & \cdots & -x \\
\vdots & & \vdots & & \vdots \\
-x & \cdots & w & \cdots & -x \\
\vdots & & \vdots & & \vdots \\
-x & \cdots & -x & \cdots & w
\end{bmatrix}
\end{array}
$$

where w is the gain associated with a correct prediction, and x is the loss associated with an incorrect prediction. We shall henceforth refer to this case as the symmetric payoff condition. We will later consider modifications of the axioms appropriate to the more general nonsymmetric case in which the amount of gain or loss varies as a function of the particular response-event combinations.

Stimulus Axiom. The stimulus situation associated with the onset of each trial is represented by a set of N stimulus elements. On each trial exactly one element is randomly sampled from this set.

Conditioning-State Axiom. On every trial each stimulus element is conditioned to exactly one response; furthermore, the element is either strongly or weakly conditioned to that response. (The strong conditioning state for the A_i response is denoted by S_i, the weak state by W_i.)

Response Axiom. If the sampled element is conditioned to the A_i response (either weakly or strongly) then that response will occur with probability 1.

Conditioning Axioms.

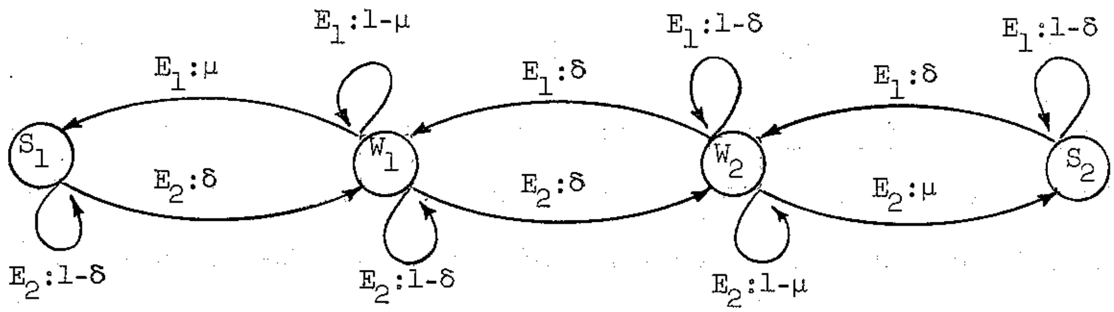
C1. Stimulus elements that are not sampled on a trial do not change their conditioning state.

C2. If event E_i occurs, then (a) if the sampled element is strongly conditioned to the A_i response it reamins so and (b) if the sampled element is weakly conditioned to the A_i response there is a probability μ that it becomes strongly conditioned.

C3. If event E_j occurs $(i \neq j)$, then (a) if the sampled element is strongly conditioned to the A_i response there is a probability δ

that it becomes weakly conditioned to A_i and (b) if the sampled element is weakly conditioned to the A_i response there is a probability δ that it becomes weakly conditioned to the A_j response.

Figure 1 illustrates the transitions that are possible under the conditions of Axioms C2 and C3 for the two-response case. Note that an element can transit only to a directly adjoining conditioning state.



Mathematical Development

Asymptotic choice proportions in the two-response case.

We begin by considering the two-response case in which the matrix of outcomes is symmetric. The event probabilities are specified by π_i, the probability of event E_1 on trial n, given that response A_i was made on that trial. Thus

$$(1) \qquad \begin{aligned} \pi_1 &= P(E_{1,n}|A_{1,n}) & \pi_2 &= P(E_{1,n}|A_{2,n}) \\ 1-\pi_1 &= P(E_{2,n}|A_{1,n}) & 1-\pi_2 &= P(E_{2,n}|A_{2,n}) \end{aligned}$$

Now assume that the k^{th} element is sampled on some trial n. The tree diagrams of Fig. 2 illustrate how the conditioning states of that element may change. For example, suppose that the sampled element is in state S_1. By the Response Axiom, the subject will make an A_1 response, which will be reinforced with probability π_1 and not reinforced with probability $1-\pi_1$. By Axiom C2, if the response is reinforced, the conditioning state will not change. If the response is not reinforced then (by Axiom C3) with probability δ the conditioning state becomes W_1. Similar applications of the axioms permit us to completely specify the possible ways in which each conditioning state may change.

We will denote the subsequence of trials on which the k^{th} stimulus element is sampled by ω_k. We next define a random variable associated with the k^{th} element that takes the conditioning states S_1, W_1, W_2, and S_2 as its values. It can be shown that over the subsequence of trials, ω_k, the random variable forms a Markov chain.

From Fig. 2, we may derive the following transition matrix for the Markov chain.

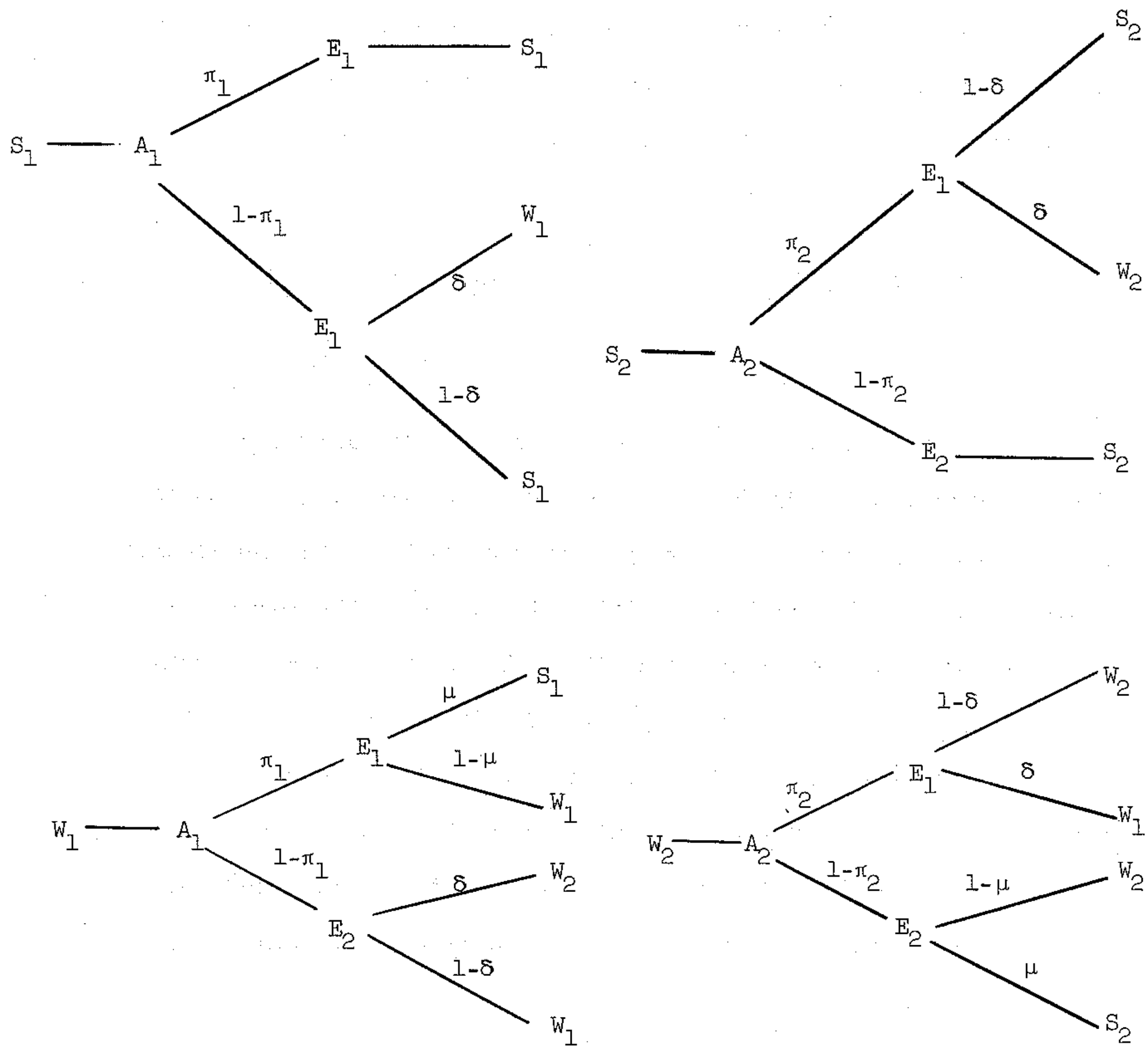


Figure 2. Transitions among conditioning states for the subset of trials on which an element is sampled.

$$
(2) \quad
\begin{array}{c} \\ S_1 \\ W_1 \\ W_2 \\ S_2 \end{array}
\begin{array}{c}
\begin{array}{cccc} S_1 & W_1 & W_2 & S_2 \end{array} \\
\begin{bmatrix}
1-\delta(1-\pi_1) & \delta(1-\pi_1) & 0 & 0 \\
\mu\pi_1 & 1-\mu\pi_1-\delta(1-\pi_1) & \delta(1-\pi_1) & 0 \\
0 & \delta\pi_2 & 1-\delta\pi_2-\mu(1-\pi_2) & \mu(1-\pi_2) \\
0 & 0 & \delta\pi_2 & 1-\delta\pi_2
\end{bmatrix}
\end{array} .
$$

For simplicity the states will be numbered as follows: $1 = S_1$, $2 = W_1$, $3 = W_2$ and $4 = \delta_2$. Next, consider the quantity $p_{ij}^{(k,m)}$, the probability of the k^{th} element being in state j on the m^{th} trial of subsequence ω_k, given that on trial 1 the element was in state i. Since the four-state Markov chain defined by Eq. 2 is irreducible and aperiodic, the quantity u_j exists, where

$$
(3) \qquad u_j = \lim_{m\to\infty} p_{ij}^{(m)} .
$$

The u_j's may be computed by

$$
u_j = \frac{D_j}{D_1+D_2+D_3+D_4} ,
$$

where

$$
(5) \qquad D_1 = \pi_1\pi_2^2 , \qquad D_3 = (1-\pi_1)^2\pi_2\,\varphi ,
$$

$$
D_2 = (1-\pi_1)\pi_2^2\,\varphi , \qquad D_4 = (1-\pi_2)(1-\pi_1)^2 ,
$$

and $\varphi = \delta/\mu$.

Atkinson (1962) has shown that at asymptote, the probability of an A_1 response for the complete N-element process is a simple function of the u_j's. Specifically,

(6) $\lim_{n\to\infty} P(A_{1,n}) = u_1 + u_2$

$$= \frac{\pi_1\pi_2^2 + (1-\pi_1)\pi_2^2\varphi}{\pi_1\pi_2^2 + (1-\pi_2)(1-\pi_1)^2 + [(1-\pi_1)\pi_2^2 + (1-\pi_1)^2\pi_2]\varphi} \ .$$

Note that the expression is independent of N, the number of stimulus elements. Henceforth, to simplify notation the trial subscripts will be omitted when asymptotic expressions are referred to; i.e.,

$$\lim_{n\to\infty} P(A_{1,n}) = P(A_1) \ .$$

Experiments in which noncontingent reinforcement is employed are frequently encountered; reinforcement is said to be noncontingent if the occurrence of an E_i event on trial n is independent of the response made on that trial. In such instances $\pi_1 = \pi_2 = \pi$, and Eq. 6 simplifies to

(7) $$P(A_1) = \frac{\pi^3 + \pi^2(1-\pi)\varphi}{\pi^3 + (1-\pi)^3 + \pi(1-\pi)\varphi} \ .$$

If

(8) $$0 \leq \mu \leq 1 \ , \quad 0 < \delta \leq 1 \ ,$$

then $P(A_1)$ is a monotonically decreasing function of φ and has the bounds

(9) $$\pi \leq P(A_1) < \frac{\pi^3}{\pi^3 + (1-\pi)^3} \ .$$

If μ equals 0, then the transition matrix for each element is reduced to two states, W_1 and W_2. In this case we have a one-stage, N-element model with a single conditioning parameter δ, and the limit of $P(A_{1,n})$ is π. This special case of our model is precisely that described by Estes (1959) as the "pattern" model. The more general formulation presented in this paper has the advantage over the pattern model of being able to account for observed values of $P(A_1)$ greater than π. Such values are generally obtained in choice experiments involving animals, choice experiments involving human subjects playing for monetary payoffs, and frequently in human choice experiments not involving payoff when run for several hundred trials.

When δ equals μ then

$$P(A_1) = \frac{\pi^2}{\pi^2 + (1-\pi)^2} \quad . \tag{10}$$

This last result is of special interest, since it is predicted by the "scanning" model developed by Estes (1962), and has been shown to give a fairly good account of several sets of data obtained in monetary payoff experiments.

<u>Sequential statistics for the noncontingent two-response case.</u>

Statistics that reflect the sequences of responses and events are of special interest. Such statistics may discriminate among models when the statistic $P(A_1)$ does not; furthermore, for our model they provide a basis for the estimation of the full array of parameters. In this paper, we will apply the model to first-order conditional probabilities of the form

$$P(A_{i,n+1} | E_{j,n} A_{k,n}) ,$$

and, when the number of observations permits, to second-order conditional probabilities of the form

$$P(A_{i,n+1} | E_{j,n} A_{k,n} E_{\ell,n-1} A_{m,n-1}) ,$$

for $i, j, k, \ell, m, = 1,2$. The presentation will be restricted to asymptotic predictions for the noncontingent two-response case, though extensions to more complex situations and to preasymptotic data can be obtained. In referring to the asymptotic statistics, we will drop the trial subscripts with the order in time from right to left being understood; thus

$$(11) \qquad \lim_{n \to \infty} P(A_{i,n+1} | E_{j,n} A_{k,n}) = P(A_i | E_j A_k)$$

$$\lim_{n \to \infty} P(A_{i,n+1} | E_{j,n} A_{k,n} E_{\ell,n-1} A_{m,n-1}) = P(A_i | E_j A_k E_\ell A_m) .$$

The derivation of these statistics is lengthy and will not be presented here. However, the general approach is developed in several sources, notably Suppes and Atkinson (1960) and Atkinson and Estes (1963). The first-order conditional probabilities are as follows:

$$(12) \qquad \begin{aligned} P(A_1 | E_1 A_1) &= \frac{N-1}{N} (u_1 + u_2) + \frac{1}{N} , \\ P(A_1 | E_2 A_1) &= \frac{N-1}{N} (u_1 + u_2) + \frac{1}{N} \left[\frac{u_1 + u_2(1-c)}{u_1 + u_2}\right] , \\ P(A_1 | E_1 A_2) &= \frac{N-1}{N} (u_1 + u_2) + \frac{1}{N} \left(\frac{u_3 c}{u_1 + u_2}\right) , \\ P(A_1 | E_2 A_2) &= \frac{N-1}{N} (u_1 + u_2) , \end{aligned}$$

where u_i is defined by Eq. 4. The next set of equations are representative of the second-order conditional probabilities. (The complete set of equations appear in Appendix A.)

$$P(A_1|E_1A_1E_1A_1) = \frac{1+(N-1)(u_1+u_2)[3+(N-2)(u_1+u_2)]}{N(N-1)(u_1+u_2)+1]} ,$$

$$P(A_1|E_1A_2E_1A_1) = \frac{(u_3+u_4)[1+(N-3)(u_1+u_2)]+u_3}{N(u_3+u_4)} ,$$

(13) $$P(A_1|E_2A_1E_1A_1) =$$

$$\frac{(u_1+u_2) - u_2\delta(1-\mu)+(N-1)(u_1+u_2)\{(u_1+u_2)[3+(N-2)(u_1+u_2)]-u_2\delta\}}{N[(N-1)(u_1+u_2)+1]}$$

$$P(A_1|E_2A_2E_1A_1) = \frac{1+(N-2)(u_3+u_4)}{N} .$$

Note that predictions of $P(A_1)$ depend only on estimates of φ; the predictions of the conditional statistics require estimates of N, δ, μ.

<u>Estimation of parameters and evaluation</u>.

The estimates of N, δ, and μ that we shall use to make predictions for the first-order conditional probabilities are those that yield a minimum value for the function

(14) $$\chi^2 = \sum_{i,j,k} \frac{n_{jk}[\hat{P}(A_i|E_jA_k) - P(A_i|E_jA_k)]^2}{P(A_i|E_jA_k)} .$$

In this equation $\hat{P}(A_i | E_j A_k)$ is the observed asymptotic conditional probability, $P(A_i | E_j A_k)$ is a predicted conditional probability based on Eq. 12, and is a function of N, μ and δ, and n_{jk} is the observed number of $E_j A_k$ occurrences, i.e., the denominator of the observed conditional probabilities. The minimum χ^2 estimates cannot be obtained analytically; however, a high-speed computer can be used to scan a grid of possible values, until estimates of N, μ and δ are obtained that minimize χ^2 to the desired degree of accuracy.

If the theory postulated that the probability distribution on trial n depends only upon the responses and reinforcing events of trial $n-1$, then the statistic described by Eq. 14 would indeed be distributed as χ^2 (Anderson and Goodman, 1957). Under these conditions the statistic would serve as a rigorous test of the ability of the model to describe the conditional sequential data. Furthermore, the estimates of the parameters would have several desirable properties common to minimum χ^2 estimates. Such estimates are consistent (as the sample size increases the estimates converge stochastically to the parameter) and asymptotically efficient (as the sample size increases, the variance of the estimates approaches the minimal variance attainable for any consistent estimate of the parameter, and the distribution of the estimate approaches the normal distribution). In the model that we have proposed, the distribution on trial n in fact depends upon the responses and reinforcement events of all preceding trials. Therefore, the statistic defined by Eq. 14 is only approximately χ^2-distributed. Its validity as a test of goodness-of-fit is not absolute, and we cannot be certain of the properties of our estimates. However, this "pseudo-χ^2" is useful as a rough index of the fit of the model, and as a means of discriminating among alternative models. Furthermore, the proximity to the χ^2 distribution improves rapidly as the number of trial outcomes upon which we are conditioning increases. With the above qualifications in mind, we will continue to refer to the statistic of Eq. 14, and to similar statistics, as χ^2. In assessing the significance level of the χ^2 of Eq. 14, we shall assume that it is distributed on one degree of freedom (df) when based on the data of a single experimental group. There are initially 8 df, one for each observed conditional probability. However, only four of these observations are independently distributed since $P(A_1|E_jA_k)+P(A_2|E_jA_k) = 1$. An additional degree of freedom is then subtracted for each parameter estimate, leaving 1 df.

The minimum χ^2 procedure may also be used to obtain parameter estimates from the second-order conditional data. In this case, it is necessary to obtain the set of estimates that minimize

$$(15) \qquad \chi^2 = \sum_{i,j,k,\ell,m} \frac{n_{jk\ell m}[P(A_i|E_jA_kE_\ell A_m) - P(A_i|E_jA_kE_\ell A_m)]^2}{P(A_i|E_jA_kE_\ell A_m)} .$$

Here $n_{jk\ell m}$ is the observed number of $E_jA_kE_\ell A_m$ occurrences. As a test of the data from a single experimental group this χ^2 would be distributed on 13 df; i.e., there are initially $2 \times 2 \times 2 \times 2 = 16$ degrees of freedom but three parameters are estimated from the data and χ^2 is therefore interpreted with $16-3 = 13$ df.

Data Analyses in the Two-Response Case

Suppes and Atkinson study.

Suppes and Atkinson (1960, ch. 10) ran 3 groups of 30 subjects each for 240 trials in a noncontingent two-choice situation. The groups differed with respect to the amount of payoff. Group Z had no monetary gains or losses; Group F gained 5¢ for each correct response and lost 5¢ for each incorrect response; and Group T gained or lost 10¢. The values of π was .6 for all groups. Table 1 contains predicted and observed values of $P(A_1)$ and of the first-order conditional probabilities, the minimum χ^2 estimates of N, δ and μ, and the values of the minimum χ^2. All observations are based on the last block of 80 trials.[2] The model describes the conditional probabilities exceedingly well, whether we merely compare the observed and predicted values or look at the values of the minimum χ^2. The mean absolute difference between observed and predicted values is approximately .005 and the sum of χ^2 over the three experimental groups is 1.85, which is not significant at even the 50% level with 3df.

The parameter estimates indicate that, under zero payoff, an incorrect response is far more likely to result in a change of conditioning state than is a correct response. Since μ exhibits a more rapid increase than δ does over groups, the discrepancy between the effects of gains and losses decreases as payoff increases. Further, the decrease in N, the number of stimulus elements, with increased payoff suggests that the subjects attend to fewer cues as motivation is increased. These conclusions, that follow from the values of the estimated parameters, suggest how the model may provide an interpretation of the effects of gains and losses.

Table 1

Observed and Predicted Values, Parameter Estimates and Minimum χ^2 for the Suppes and Atkinson Experiment

(Observed values are given in parentheses)

	Group		
	Z	F	T
$P(A_1)$	.605 (.600)	.648 (.649)	.695 (.700)
$P(A_1 \mid E_1A_1)$	.707 (.709)	.792 (.794)	.855 (.855)
$P(A_1 \mid E_2A_1)$	.534 (.534)	.603 (.601)	.660 (.670)
$P(A_1 \mid E_1A_2)$	.626 (.606)	.615 (.613)	.624 (.638)
$P(A_1 \mid E_2A_2)$	.450 (.449)	.382 (.388)	.329 (.323)
χ^2	1.08	.09	.68
μ	.02	.23	1.00
δ	.70	.69	.95
N	3.90	2.44	1.90

Myers, Fort, Katz, and Sydam's study.

Myers et at (1963) ran 9 groups of 20 subjects for 400 trials in a noncontingent two-choice experiment. Levels of π of .6, .7, and .8 were employed and subjects gained or lost 0¢, 1¢ or 10¢. The data analysis reported here is based on trials 301-400. (Due to an error in recording during one experimental session, the data of only 16 subjects are available for the .6-0¢ group.)

The minimum χ^2 procedure described in the previous section was applied to the first-order conditional probabilities of the Myers et al study with one modification. The estimates reported for each payoff level in Table 2 are those that minimized a sum of χ^2 over the three levels of π. Thus there are 9 df associated with each χ^2 value in Table 2; i.e., twelve predictions were made on the basis of three parameters at each payoff level. As in the Suppes and Atkinson study, μ increases more rapidly than does δ and N decreases with increasing payoff; unlike the results of that study, the statistics for the 1¢ and 10¢ groups are not widely disparate, and this is reflected in the closeness of the estimates for those groups. The minimum χ^2 are fairly large and it appears that the model does not adequately describe the data. We will consider this point further when we look at the actual observations and predictions.

Table 3 presents observed and predicted values of $P(A_1)$, obtained by inserting the minimum estimates of δ and μ into Eq. 7. With the exception of the .6-1¢ and .7-1¢ groups, the fit appears quite good; the overall mean deviation of observed and predicted values is 1.6%. At least for the 0¢ and 10¢ groups, single values of φ give a

Table 2

Values of the Parameter and Minimum Chi-Squares for the Myers et at Study

	Payoffs		
	0¢	1¢	10¢
N	5.05	2.21	2.19
δ	.82	1.00	1.00
μ	.13	.83	1.00
χ^2	52.27	48.06	24.98

Table 3

Predicted and Observed Values of $P(A_1)$ for the Myers et al Study

(Observed values are given in parentheses)

	π		
Payoffs	.6	.7	.8
0¢	.627	.749	.863
	(.624)	(.753)	(.869)
1¢	.684	.835	.934
	(.653)	(.871)	(.925)
10¢	.692	.845	.941
	(.714)	(.866)	(.951)

reasonably good account of the data at three different levels of π. Despite the large X^2 values reported above, this is an important result. That adequate predictions of $P(A_1)$ under payoff is not trivial is suggested by the fact that several theories (Edwards, 1956; Siegel, 1959; Estes, 1962) have been developed for this purpose alone.[3]

Table 4 presents the observed and predicted first-order conditional probabilities. The most notable aspect of the table is the fact that, excluding the .6-0¢ group, statistics of the form $P(A_1|E_jA_1)$ are accurately predicted; the major source of the large values of X^2 appears to be the failure to predict statistics of the form $P(A_1|E_jA_2)$. This conclusion is supported by the fact that a X^2 computed for the 16 E_jA_1 statistics is 13.23 which is not significant at the .05 level on 7 df. There are at least two obvious explanations for the poor fit of the E_jA_2 statistics: (a) it is possible that these statistics are unreliable since they are generally based on fewer observations than the E_jA_1 statistics (this argument might also be applied to the data of the .6-0¢ group which are based on 20% fewer observations than those of the other groups); (b) the model may require some modification to adequately describe the statistics that are poorly fit in Table 4. One argument against the second conclusion is the fact that the relationship between observations and predictions is not consistent; if the defect was in the theory rather than in the data, the predictions might be expected to be consistently too high, or consistently too low. However, additional experimentation involving more trials and subjects is required to decide between these two alternatives.

Table 4

Predicted and Observed Values of $P(A_1|E_iA_j)$ for the Myers et al Study

(Observed values are given in parentheses)

Payoffs	π	$P(A_1\|E_1A_1)$	$P(A_1\|E_2A_1)$	$P(A_1\|A_1A_2)$	$P(A_1\|E_2A_2)$
0¢	.6	.701 (.668)	.569 (.484)	.649 (.726)	.503 (.593)
	.7	.799 (.816)	.680 (.666)	.753 (.747)	.601 (.571)
	.8	.890 (.901)	.790 (.824)	.848 (.746)	.692 (.803)
1¢	.6	.810 (.818)	.614 (.609)	.643 (.613)	.356 (.336)
	.7	.930 (.939)	.780 (.789)	.806 (.825)	.475 (.415)
	.8	.959 (.947)	.858 (.856)	.878 (.974)	.505 (.588)
10¢	.6	.837 (.849)	.655 (.674)	.674 (.653)	.406 (.407)
	.7	.924 (.923)	.786 (.786)	.803 (.865)	.492 (.471)
	.8	.972 (.974)	.879 (.873)	.892 (.906)	.540 (.923)

Friedman et al study.

Friedman et al (1963) ran 80 subjects for three sessions of 384 trials each in a noncontingent two-choice experiment. No monetary payoff was involved. During the first two sessions, π was varied among blocks of 48 trials. In the third session (following 48 trials at a π value of .5) subjects were tested for 288 trials at a π value of .8. The analyses of tables 5 and 6 are based on trials 193-288 of the .8 series. Parameter estimates were obtained from the first-order statistics by minimizing the χ^2 of Eq. 14. These estimates are the basis for both the predictions of the first-order conditional probabilities of Table 5 and the second-order conditional probabilities of Table 6. We have also investigated the possibility of obtaining parameter estimates from the second-order data by minimizing the χ^2 of Eq. 15; the results of the two procedures differ very little, and consequently we present only the results based on parameter estimates for the first-order data.

The first-order conditional statistics are fairly well fit; the second-order statistics appear to present a problem. The χ^2 defined by Eq. 15 is 60.48 on 13 df and several predictions clearly deviate from the observations. The fit is particularly poor for those statistics that are based on the fewest observations whereas the description of the more reliably based $P(A_1|E_1A_1E_1A_m)$ is quite reasonable. More experiments providing large numbers of observations are required before we can conclude that the model fails to predict higher-order conditional statistics. However, these results, together with other data (Anderson, 1963), suggest that prediction of response probabilities conditioned on the outcomes of several trials is a major problem for models of this type.

Table 5

Predicted and Observed Values of $P(A_1|E_jA_k)$, Parameter Estimates, and minimum χ^2 Value for the Friedman et al Study

	Observed	Predicted
$P(A_1\|E_1A_1)$	.894	.899
$P(A_1\|E_2A_1)$	.744	.730
$P(A_1\|E_1A_2)$	.692	.693
$P(A_1\|E_2A_2)$	.407	.489
δ	.50	
μ	.03	
N	2.44	
χ^δ	9.68	

Table 6

Observed and Predicted Values of $P(A_1|E_jA_kE_lA_m)$

for the Friedman et al Study

	Observed	Predicted
$P(A_1\|E_1A_1E_1A_1)$	.925	.937
$P(A_1\|E_1A_1E_1A_2)$	.817	.803
$P(A_1\|E_1A_1E_2A_1)$	.848	.833
$P(A_1\|E_1A_1E_2A_2)$	.610	.559
$P(A_1\|E_1A_2E_1A_1)$	.747	.763
$P(A_1\|E_1A_2E_1A_2)$	.606	.648
$P(A_1\|E_1A_2E_2A_1)$	.769	.657
$P(A_1\|E_1A_2E_2A_2)$	.523	.621
$P(A_1\|E_2A_1E_1A_1)$	.801	.770
$P(A_1\|E_2A_1E_1A_2)$	.603	.623
$P(A_1\|E_2A_1E_2A_1)$	.595	.662
$P(A_1\|E_2A_1E_2A_2)$	.519	.390
$P(A_1\|E_2A_2E_1A_1)$	.600	.559
$P(A_1\|E_2A_2E_1A_2)$	.483	.444
$P(A_1\|E_2A_2E_2A_1)$	.257	.452
$P(A_1\|E_2A_1E_2A_2)$	.220	.421

A contingent reinforcement study.

Thus far, all the studies considered in this section have involved a noncontingent reinforcement procedure. Experiments using the contingent reinforcement procedure are relatively rare, and we know only one such study in which monetary payoff was involved. Since only three values of $P(A_1)$ are involved, our analysis hardly constitutes a test of the model for contingent experiments. However, the results are encouraging. Atkinson (1962) ran 3 groups of 20 subjects each for 340 trials, with each correct response resulting in a gain of 5¢ and each incorrect response resulting in a loss of 5¢. The groups differed with respect to π_1 which took the values .6, .7, and .8; for all groups π_2 equalled .5. Table 7 presents the observed proportions of A_1 responses for the last 80 trials. The predicted values were obtained by inserting a least-squares estimate of φ into Eq. 6. The estimate $\varphi = 2.1$, results in a good acount of the values of $P(A_1)$; the mean absolute deviation of observed from predicted is about 1%.

Analyses of the Three-Response Case

We will now consider an extension of the model to experiments involving three responses. Since the only available data have been obtained for noncontingent procedures, equations will be presented only for that case. However, a more general statement is easily obtained following the approach of the previous section. For the noncontingent case, the axioms presented earlier result in the following transition matrix for element k over a subsequence w_k of trials.

Table 7

Observed and Predicted Values of $P(A_1)$ for

a Contingent Reinforcement Experiment

π_1	Observed	Predicted
.6	.601	.592
.7	.685	.704
.8	.832	.831

$$
(16)\qquad
\begin{array}{c}
\\
S_1\\ S_2\\ S_3\\ W_1\\ W_2\\ W_3
\end{array}
\begin{array}{c}
\begin{array}{cccccc}
S_1 & S_2 & S_3 & W_1 & W_2 & W_3
\end{array}\\
\begin{bmatrix}
1-\delta(1-\gamma_1) & 0 & 0 & \delta(1-\gamma_1) & 0 & 0\\
0 & 1-\delta(1-\gamma_2) & 0 & 0 & \delta(1-\gamma_2) & 0\\
0 & 0 & 1-\delta(1-\gamma_3) & 0 & 0 & \delta(1-\gamma_3)\\
\mu\gamma_1 & 0 & 0 & 1-\mu\gamma_1-\delta(1-\gamma_1) & \delta\gamma_2 & \delta\gamma_3\\
0 & \mu\gamma_2 & 0 & \delta\gamma_1 & 1-\mu\gamma_2-\delta(1-\gamma_2) & \delta\gamma_3\\
0 & 0 & \mu\gamma_3 & \delta\gamma_1 & \delta\gamma_2 & 1-\mu\gamma_3-\delta(1-\gamma_3)
\end{bmatrix}
\end{array}.
$$

For this case we let γ_i denote the probability of event E_i, where $\gamma_1 + \gamma_2 + \gamma_3 = 1$. The states will be designated by numbers corresponding to the ordering in the matrix, i.e., $S_1 = 1$, etc. From Eq. 16 we obtain the u_j defined previously by Eq. 3:

$$
(17)\qquad u_j = \frac{D_j}{D_1+D_2+D_3+D_4+D_5+D_6} ,
$$

where

$$
(18)\qquad
\begin{aligned}
D_1 &= \gamma_1^2(1-\gamma_2)(1-\gamma_3) , & D_4 &= \lambda_1(1-\lambda_1)(1-\gamma_2)(1-\gamma_3)\varphi ,\\
D_2 &= \gamma_2^2(1-\gamma_1)(1-\gamma_3) , & D_5 &= \gamma_2(1-\gamma_1)(1-\gamma_2)(1-\gamma_3)\varphi ,\\
D_3 &= \gamma_3^2(1-\gamma_1)(1-\gamma_2) , & D_6 &= \gamma_3(1-\gamma_1)(1-\gamma_2)(1-\gamma_3)\varphi ,
\end{aligned}
$$

and again $\varphi = \delta/\mu$. Following the procedure for the two-response case we may derive expressions for $\lim_{n\to\infty} P(A_{i,n}) = P(A_i)$, specifically

$$P(A_1) = u_1 + u_4 ,$$

(19)
$$P(A_2) = u_2 + u_5 ,$$

$$P(A_3) = u_3 + u_6 .$$

Cotton and Rechtshaffen's study.

Cotton and Rechtshaffen (1958) report values of $P(A_1)$ for six groups, two having two responses available, and four having three responses available. Values of $P(A_1)$ and standard deviations of proportions for trials 286-450 are presented in Table 8, together with predictions derived from the model. A least-squares estimation procedure yielded a value of φ of 3.7; this was then substituted into Eq. 19 with the appropriate γ_i values to generate the six predictions. The average absolute deviation of observed from predicted values is less than 1.4%, which is quite small in view of the variability in the proportions. It is particularly interesting to note that the finding that $P(A_1)$ increases as the number of choices increases (see also Gardner, 1957) is accounted for by the present model.

Cole's study.

Cole (1962) ran three groups of human subjects under a noncontingent reinforcement procedure. Two of the groups had three responses available, with γ_i's of 2/3, 2/9, and 1/9 for one group, and γ_i's of 4/9, 1/3, and 2/9 for the second group; the third group had two responses available with γ equal to 2/3. Table 9 presents the observed and predicted values of $P(A_1)$ and values of φ for each group and response. The values of φ were computed by solving Eq. 19, and the observations were based on trials 501-1000.

Table 8

Observed and predicted $P(A_1)$ for the Cotton Rechtschaffen experiment

Condition	Predicted $P(A_1)$	Observed $P(A_1)$	s_p
60-40	.641	.614	.118
60-30-10	.658	.658	.112
60-20-20	.671	.660	.096
70-30	.773	.741	.099
70-20-10	.783	.801	.137
70-15-15	.784	.805	.091

Table 9

Observed and predicted $P(A_i)$ for the Cole experiment

Condition	Response	Predicted $P(A_i)$	Observed $P(A_i)$	φ
$\frac{2}{3}-\frac{2}{9}-\frac{1}{9}$	A_1	.844	.881	.45
	A_2	.109	.087	.43
	A_3	.047	.029	.41
$\frac{4}{9}-\frac{1}{3}-\frac{2}{9}$	A_1	.512	.531	.52
	A_2	.313	.304	.40
	A_3	.174	.165	.55
$\frac{2}{3}-\frac{1}{3}$	A_1	.812	.779	1.50
	A_2	.188	.221	1.50

Averaging over responses and then over groups, a value of φ of .81 was obtained; substitution in Eq. 19 resulted in the predicted $P(A_i)$ of Table 9. The model again correctly predicts an increase in the value of $P(A_1)$ as the number of response alternatives increases from two to three, but the discrepancies between observations and predictions are somewhat greater than they were for the Cotton and Rechtshaffen data. This is due to the difference in the average values of φ for two and three choice data. If predictions are made just for the data from the three-choice groups (using a value of φ based only on the observations for those groups) the average difference between the observed and predicted values of $P(A_i)$ is only .5%. It is possible that different values of φ are required for each number of response alternatives. However, the fit for the Cotton and Rechtshaffen data could argue against this assumption. Additional experimentation involving varying numbers of response alternatives is required for clarification of this problem.

Extension of the Model to Response Times

Despite recurrent attempts to develop an adequate theory of response times (Estes, 1951; Bush and Mosteller, 1955, LaBerge, 1959; Luce, 1960) this dependent variable has proven more elusive than response probability. One attractive feature of the weak-strong model is that it can be extended to treat response times with the addition of only one assumption. Furthermore, derivations of a variety of statistics are extremely simple, and estimates of response time parameters can be easily obtained. To facilitate the presentation, we will limit the discussion to the asymptotic case for the two-response noncontingent situation. Extensions to more complex situations and to preasymptotic data follow readily from the developments of this section.

The set of axioms previously presented for choice behavior are still assumed to hold. Thus, we postulate that exactly one element is sampled on each trial, that the element is either weakly or strongly conditioned to one of the response alternatives, and that the conditioning state may change in accord with the previously presented conditioning axioms. In addition, we require the following axiom:

<u>Response Time Axiom</u>. The random variable T_n denotes the response time on trial n of the experiment and depends on the conditioning state of the sampled stimulus element. If the sampled element is in a strong state of conditioning, then the distribution of response times has probability density $S(t)$ with finite mean s. If the sampled element is in a weak state of conditioning, then the distribution of response times has probability density $W(t)$ with finite mean w.

On the basis of the response time axiom and our choice model, a number of predictions may be derived. We will next consider some of these. Since all equations will be for the asymptotic case, the subscript n with be omitted.

<u>Mean Response Times</u>.

$E(T)$, the mean asymptotic response time obtained by averaging over both A_1 and A_2 responses is simply the weighted sum of s and w, where the weights are the probabilities of sampling from the two hypothesized distributions. Accordingly, we have

$$E(T) = (u_1+u_4)s+(u_2+u_3)w$$

(20)

$$= \frac{s[\pi^3+(1-\pi)^3] + w[\pi(1-\pi)\varphi]}{\pi^3+(1-\pi)^3 + \pi(1-\pi)\varphi} .$$

If we assume that $s < w$, which appears reasonable, then it is easily proven that the mean response time is greatest when $\pi = .5$, and monotonically decreases as π approaches one.

We next consider $E(T|A_i)$, the mean response time for an A_i response. This quantity is derived as the weighted sum of s and w, where the weights are the probabilities of sampling from the two hypothesized distributions, given that an A_i response has occurred. The appropriate equations are

$$E(T|A_1) = \frac{su_1+wu_2}{u_1+u_2}$$

(21)

$$= \frac{s\pi+w(1-\pi)\varphi}{\pi+(1-\pi)\varphi} .$$

(22)

$$E(T|A_2) = \frac{s(1-\pi)+w\pi\varphi}{1-\pi+\pi\varphi} .$$

Once φ has been estimated from the choice data, the parameters s and w may be simultaneously solved for in Eqs. 21 and 22. Predictions of $E(T)$, $E(T|A_1)$, and $E(T|A_2)$ can then be made for any value of π.

If $s < w$, it can be shown that the mean time required for a response to occur is a monotone decreasing function of the probability of the predicted event. Response time data from the Friedman et al study are ambiguous with regard to this prediction. Response times for the A_1 response were slightly (but significantly) less than A_2

response times, as predicted; however, response time did not vary as a function of π.[4] Data that are more clearly consistent with the predictions of this model are reported by Calfee (1963) who found that response times for rats decreased as π increased, and that the preferred response was made more quickly than the less preferred response. These data support the weak-strong model, and suggest that LaBerge's (1959) neutral elements model requires revision. That model predicts no differences in average A_1 and A_2 response times, or in response times as a function of π.

We conclude this section by presenting equations for statistics of the form $E(T|A_iE_jA_k)$, the expected response time of an A_i response on trial $n+1$, given that it was preceded by event E_j and response A_k on trial n. The general form of the expression for this statistic is

$$E(T|A_iE_jA_k) = \frac{sP(S_i|E_jA_k) + wP(W_i|E_jA_k)}{P(A_i|E_jA_k)} \quad (23)$$

In the above expression $P(S_i|E_jA_k)$ denotes the asymptotic probability that an element is strongly conditioned to A_i on trial $n+1$ given that E_jA_k occurred on trial n; $P(W_i|E_jA_k)$ has a similar interpretation. Substituting in Eq. 23, we obtain the following expressions:

$$E(T|A_1E_1A_1) = \frac{(N-1)(su_1+wu_2)+(\frac{1}{u_1+u_2})[s(u_1+u_2\mu)+wu_2(1-\mu)]}{(N-1)(u_1+u_2)+1}$$

$$E(T|A_1E_2A_1) = \frac{(N-1)(su_1+wu_2)+(\frac{1}{u_1+u_2})\{su_1(1-\delta)+w[u_1\delta+u_2(1-\delta)]\}}{(N-1)(u_1+u_2)+(\frac{1}{u_1+u_2})[u_1+u_2(1-\delta)]}$$

(24)

$$E(T|A_1E_1A_2) = \frac{(N-1)(su_1+wu_2)+\frac{wu_3\delta}{u_3+u_4}}{(N-1)(u_1+u_2)+\frac{u_3\delta}{u_3+u_4}}$$

$$E(T|A_1E_2A_2) = \frac{su_1+wu_2}{u_1+u_2} \quad .$$

The expressions for the $E(T|A_2E_jA_k)$ are obtained by substituting u_4 for u_1, u_3 for u_2, and vice verse, in Eq. 24; e.g.,

$$E(T|A_2E_1A_1) = \frac{su_4+wu_3}{u_4+u_3} \; .$$

Extension of the Model to the Differential Payoff Case

Thus far we have considered a model that is applicable only to the symmetric payoff case, in which the amount gained is the same for all correct responses, and the amount lost is the same for all incorrect responses. We next consider an extension of the model to the nonsymmetric payoff case. For the two-response situation this payoff scheme may be represented by the matrix

$$\begin{array}{cc} & \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} w & -x \\ -y & z \end{bmatrix}, \end{array}$$

where the amount gained or lost is a function of the response-event combination. Although an adequate description of data obtained under such conditions would seem to be a prerequisite for a general theory of motivational variables, to date little progress has been made on the problem. Bush and Mosteller's "experimenter-subject-controlled events" model (1955, p. 286) is applicable, but this approach leads to severe mathematical difficulties. Estes' "scanning" model (1962) involves simple computations, but only yields predictions of $P(A_1)$. The same objection may be raised to Edwards' "RELM" model (1956). The generalization of the weak-strong model that we will present is mathematically tractable; the only complication beyond the original model is the need to estimate one additional parameter. The variety of predictions that follow from the original model can also be derived for the extended model. For these reasons, the generalized weak-strong model merits consideration. However, it should be noted that the developments of this section are extremely tentative. An empirical evaluation of the model is excluded at this time since there have been few experiments involving differential payoffs, and these, while theoretically suggestive, have involved too few trials and subjects to permit a test of the model.

We might extend the weak-strong model by postulating two values of μ (μ_w and μ_z), corresponding to the two gains, and the two values of δ (δ_x and δ_y); corresponding to the two losses. That this identification of parameters has limited applicability is suggested by data obtained from a matrix such as

$$\begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{bmatrix} 5 & -5 \\ 1 & 1 \end{bmatrix} \end{array} .$$

For the parameter identification proposed above, E_1 and E_2 should have identical effects upon the conditioning-state whenever the subject makes an A_2 response. Furthermore, if a 1¢ gain is assumed to be reinforcing, the subject should absorb on A_2. Both inferences are contradicted by experimental data (Myers and Sadler, 1960; Myers and Fort, 1961). A mechanism is required which permits the A_2 response to be strengthened or weakened following a 1¢ gain, depending on which event occurred.

The concept of regret (Savage, 1954) provides one approach to the problem just posed. Regret is the difference between the obtained payoff and the maximum possible payoff, given that event E_i occurs. Thus, for the last payoff matrix presented, we have the regret matrix

$$\begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{bmatrix} 0 & 4 \\ 6 & 0 \end{bmatrix} \end{array} .$$

In general, corresponding to the payoff matrix

$$\begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{bmatrix} w & x \\ y & z \end{bmatrix} \end{array} ,$$

where $y < w$ and $x < z$ (note that x and y are not necessarily negative), we have

$$\begin{array}{cc} & \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{array}{c} A_1 \\ A_2 \end{array} & \begin{bmatrix} 0 & r_1 \\ r_2 & 0 \end{bmatrix} \end{array},$$

where the regret associated with an incorrect A_1 response is

(25) $$r_1 = w - y \quad ,$$

and the regret associated with an incorrect A_2 response is

(26) $$r_2 = z - x \quad .$$

Here, we define an incorrect response as one that yields a payoff less than the maximum possible payoff, given the occurrence of E_i.

The notion of regret provides a basis for modifying the weak-strong model in the following manner. We identify μ with the probability that zero regret results in the strengthening of a correct response, δ_1 with the probability that r_1 results in the weakening of an incorrect A_1 response, and δ_2 with the probability that r_2 results in the weakening of an incorrect A_2 response. A minor change in our system of axioms now suffices in order to derive equations for choice behavior under differential payoffs, Axiom <u>C3</u> is rewritten as follows:

<u>C3'</u>. If event E_j occurs $(i \neq j)$, then (a) if the sampled element is strongly conditioned to A_i there is a probability δ_i that it becomes weakly conditioned to A_i and (b) if the sampled element is weakly conditioned to A_i there is a probability δ_i that it becomes weakly conditioned to A_j.

For the revised axioms we may now obtain the following results in the noncontingent two-choice situation:

$$(27)\qquad D_1 = \pi^2\varphi_2^2 \qquad\qquad D_3 = \pi(1-\pi)^2\varphi_2\,\varphi_1^2\ ,$$

$$D_2 = \pi^2(1-\pi)\varphi_2^2\,\varphi_1 \qquad\qquad D_4 = (1-\pi)^3\varphi_1^2\ ,$$

where

$$(28)\qquad \varphi_1 = \delta_1\mu \qquad\qquad \varphi_2 = \delta_2/\mu\ .$$

Substituting in Eq. 4 we have

$$(29)\qquad P(A_1) = \frac{\pi^3 + \pi^2(1-\pi)\varphi_1}{\pi^3+\pi^2(1-\pi)\varphi_1+\pi(1-\pi)^2\,\mathcal{E}\,\varphi_1+(1-\pi)^3\,\mathcal{E}^2}\ ,$$

where $\mathcal{E} = \varphi_1/\varphi_2$. Note that $P(A_1)$ is independent of N. Furthermore, for the symmetric case $r_1 = r_2$ and therefore $\varphi_1 = \varphi_2$; hence under this condition Eq. 29 reduces to Eq. 7. Also, the following results can be easily proved:

(i) $P(A_1)$ is bounded by zero and one. Specifically,

$$(30)\qquad \lim_{\mathcal{E}\to 0} P(A_1) = 1 \qquad \lim_{\mathcal{E}\to \infty} P(A_1) = 0\ .$$

(ii) For constant π and φ_2, $P(A_1)$ is a decreasing monotonic function of φ_1. As the regret associated with an incorrect A_1 response increases, the probability of making an A_1 decreases.

(iii) For constant π and φ_1, $P(A_1)$ is an increasing monotonic function of φ_2. As the regret associated with an incorrect A_2 response increases, the probability of making an A_1 increases.

Several experiments have recently been performed (Myers and Sadler, 1960; Myers and Katz, 1962; Katz, 1962) involving the choice between a "sure thing" and a risky option. The payoff matrices are of the form

$$\begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{bmatrix} w & -w \\ 1 & 1 \end{bmatrix} \end{array}, \quad \text{and} \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{bmatrix} w & -w \\ -1 & -1 \end{bmatrix} \end{array} .$$

Where E_1 and E_2 are equiprobable, i.e., $\pi = .5$. The major findings are that (a) $P(A_1)$ is always greater when the payoff associated with an A_2 response is -1 than when the payoff is +1, and (b) as the absolute value of w increases, $P(A_1)$ increases when the A_2 payoff is +1, and decreases when the A_2 payoff is -1. These results are schematically represented in Fig. 3. The convergence exhibited in Fig. 3 is not consistent with the results for the symmetric payoff case, in which subjects approach the optimal strategy (always predict the more frequent event) as payoffs increase. In the studies under discussion, the optimal strategy is to always make the A_1 response when the A_2 payoff is +1; subjects increasingly deviate from this strategy as the amount risked increases.

Since the convergence effect displayed in Figure 3 does not seem to be easily explained by existing theories of decision behavior, it is of interest to consider it in terms of the weak-strong model. Upon converting the above payoff matrices to regret matrices, it becomes

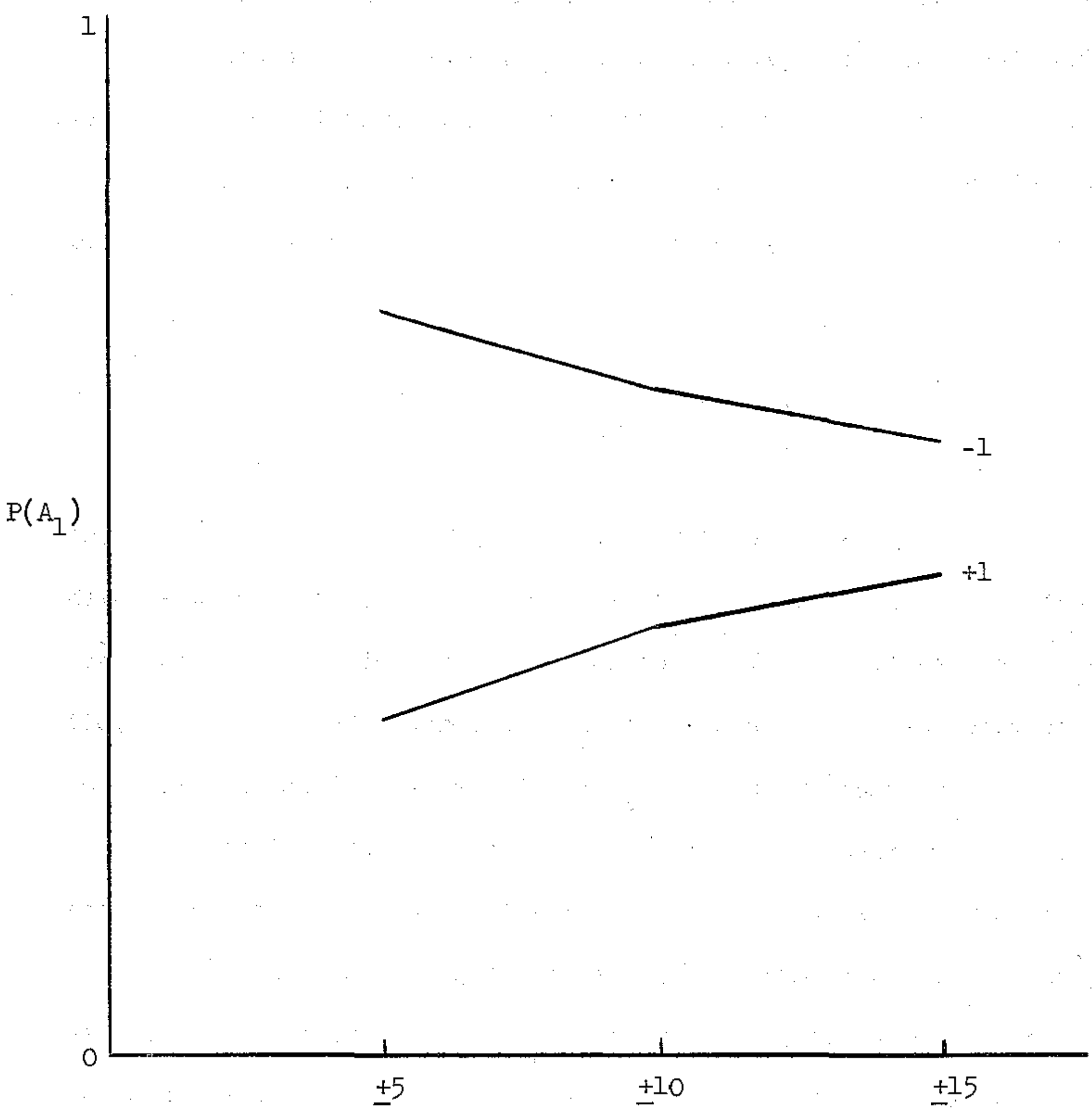


Figure 3. The proportion of risky responses as a function of the amount of risk, with the values of the "sure thing" alternative as the parameter.

apparent that r_2 is less than r_1 when the A_2 payoff is +1; and r_2 is greater than r_1 when the A_2 payoff is -1. Assuming that δ_i is a monotonically increasing function of r_i, $\mathcal{E}$ will be greater when the A_2 payoff is +1 than when it is -1. Consequently, the -1 curve should lie above the +1 curve, as it does.

We next attempt to account for the convergence depicted in Fig. 3. As w increases, both r_1 and r_2 increase, but the ratio r_1/r_2 monotonically approaches an asymptote of 1. If δ_i is a negatively accelerated function of r_i then (a) if the A_2 payoff is +1, $\mathcal{E}$ will decrease to an asymptote of 1, and (b) if the A_2 payoff is -1, $\mathcal{E}$ will increase to an asymptote of 1. Consequently, the curves should converge until they asymptote at .5. Although the above qualitative description of the risk-taking data is encouraging, an adequate evaluation of the model will require precise quantitative analyses of the data. When such analyses are available, the relationships between parameter and regret values may be more complicated than we have suggested. For example, a literal interpretation of our discussion would suggest that μ should be invariant over different payoff matrices. This is a doubtful premise, considering that such parameter invariance is often difficult to establish over levels of π. However, in view of the dearth of theories dealing with the differential payoff case, if the model even provides a reasonable account of data for a single group, some progress will have been made.

A Multi-Stage Model

As indicated earlier the "pattern" model of stimulus sampling theory would be regarded as a one-stage model. Similarly the model discussed in this paper is a two-stage model. In this section, we investigate the consequences of generalizing the model so than an element may be in one of k stages of conditioning to a response. The generalized model follows logically from the weak-strong model. The Stimulus Axiom and the Response Axiom remain unchanged; the other axioms require only the obvious modifications.

Conditioning-State Axiom. On every trial each stimulus element is conditioned to exactly one response; furthermore, the element is in one of k stages of conditioning to that response. (An element in conditioning state C_{im} is in stage m of conditioning to response A_i where $m=1,2,\ldots,k$ and k denotes the strongest stage.)

Conditioning Axioms.

C2'. If event E_i occurs, then (a) if the sampled element is in state C_{ik} it remains so and (b) if the sampled element is in state C_{im} $(m \neq k)$ there is a probability μ that it enters state $C_{i,m+1}$.

C3'. If event E_j occurs $(i \neq j)$, then (a) if the sampled element is in state C_{im} $(m \neq 1)$, there is a probability δ that it enters state $C_{i,m-1}$, and (b) if the sampled element is in state C_{i1} there is a probability δ that it enters state C_{j1}.

Figure 4 provides a schematic presentation of the transitions among states for the two-response case. It may be helpful to compare this representation with that of Fig. 1 for the weak-strong model.

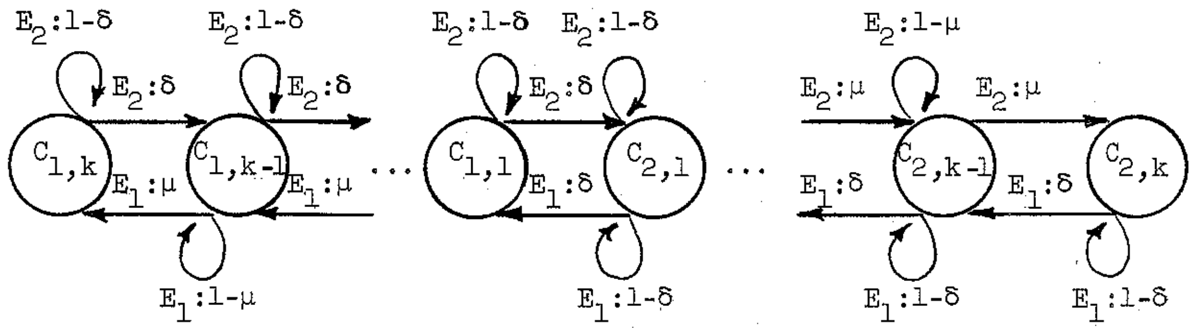


Figure 4. Possible transitions among conditioning states for the $\underline{k}$-stage model ($r = 2$).

Expressions can now be derived for those statistics previously treated in the $k = 2$ case. For the noncontingent two-response case, it can be shown that

(31) $$P(A_i) = \sum_m u_{im} ,$$

where

(32) $$u_i = \frac{D_{im}}{\sum_{i,m} D_{im}} ,$$

and

(33) $$D_{1m} = \pi^{k+m-1}[(1-\pi)\varphi]^{k-m}$$
$$D_{2m} = (\pi\varphi)^{k-m}(1-\pi)^{k+m-1} .$$

These expressions can be evaluated to yield the following equation:

(34) $$P(A_1) = \frac{\alpha}{\alpha + \beta} ,$$

where

(35) $$\alpha = \pi^k\{\pi^k + [(1-\pi)\varphi]^k\}(1-\pi-\pi\varphi)$$
$$\beta = (1-\pi)^k\{(1-\pi)^k+(\pi\varphi)^k\}[\pi-(1-\pi)\varphi] .$$

Note that

(36) $$\lim_{\varphi\to\infty} P(A_1) = \pi$$
$$\lim_{\varphi\to\infty} P(A_1) = \frac{\pi^{2k-1}}{\pi^{2k-1}+ (1-\pi)^{2k-1}} .$$

Thus, one consequence of introducing k is to increase the upper asymptotic bound on $P(A_1)$ and place it as close to 1 as desired for $\pi > 1/2$.

The form of the conditional statistics is also simple; the first-order statistics are as follows:

$$P(A_1|E_1A_1) = (\frac{N-1}{N})(\frac{\alpha}{\alpha+\beta}) + \frac{1}{N}$$

$$P(A_1|E_2A_1) = (\frac{N-1}{N})(\frac{\alpha}{\alpha+\beta}) + \frac{1}{N}\left[\frac{\alpha-u_{11}\delta(\alpha+\beta)}{\alpha}\right]$$

$$P(A_1|E_1A_2) = (\frac{N-1}{N})(\frac{\alpha}{\alpha+\beta}) + \frac{1}{N}\left[\frac{u_{21}\delta(\alpha+\beta)}{\beta}\right]$$

$$P(A_1|E_2A_2) = (\frac{N-1}{N})(\frac{\alpha}{\alpha+\beta}) .$$

We have applied the minimum χ^2 estimation procedure to the first-order conditional data presented previously for $k = 2, 3, 4, 5,$ and 10. Generally, the minimum χ^2 was smallest at $k = 2$ though there were a few instances for which it was slightly less at $k = 3$. In most instances the goodness of fit showed rapid deterioration as k increased. For example, when the models were applied to the Friedman et al data, the minimum χ^2 was at a low of 9.37 for $k = 2$, increased to 38.51 for $k = 5$, and then increased to 42,784 for $k = 10$. The increase in χ^2 appears to be due to the fact that for large k, the model predicts more response perserveration, following a correct response, than actually occurs. In view of these analyses we are prone to conclude that significant improvements in goodness of fit will not follow as a result of increases in k, and that two-stages generally will best describe the data. Assuming that this conclusion holds for future analyses of data, it, of course, applies only to our particular statement of the model. The question of k-stage models involving different response or sampling axioms remains to be investigated.

Discussion

Several articles (Bower, 1959; Atkinson, 1961; Estes, 1960, 1962) have recently demonstrated that a more molecular analysis of the subject's pre-response behavior may prove fruitful in formulating a choice model. It is therefore interesting to note that at least one such analysis of choice behavior results in the same equations derived for the weak-strong model. Specifically, consider a model which postulates that associated with each response alternative is a tendency to approach or aviod that alternative. Further assume that the set of approach tendencies, and the order in which response alternatives are considered (or observed) determine the subject's choice on any trial, and are themselves determined by the outcomes of preceding trials.

To formalize these notions let the function v_k be the approach tendency associated with response A_k. When the subject observes the kth alternative, he will make that response if $v_k = 1$, or move on to observe some other alternative if $v_k = 0$. The values of v_k for the r-response alternatives will be represented by a vector $V = < v_1, v_2, \ldots, v_r >$. For example, $V = < 001 >$ indicates that the subject will approach A_3 when he observes it; all other alternatives will be avoided. We further assume that, in the time period immediately preceding his choice, the subject orients towards each response alternative in some sequence, until he observes an alternative for which v_k is one, or, if all v_k's are zero, until he has observed each alternative. In either case, the alternative chosen is the one last observed by the subject. Thus, the subject will choose the first observed alternative for which v_k is one, or, if all values are zero, the last alternative observed. The sequence in which the alternatives are observed on any given trial will be represented by the vector $O = \langle o_1, o_2, \ldots, o_r \rangle$. The value of o_i indicates which response will be observed at position i in the observing sequence.

With these concepts in mind we can define the conditioning state of the subject on any trial n of an experiment as the vector $C_n = \langle O, V \rangle$. For example, if $C_n = \langle\langle 12 \rangle, \langle 10 \rangle\rangle$, then the subject initiates the trial by observing response A_1 and then makes that response. If $C_n = \langle\langle 12 \rangle, \langle 00 \rangle\rangle$ the subject first observes A_1, then A_2 and terminates the trial by choosing A_2 since both v_1 and v_2 equal 0.

To complete the analysis we need some rule for describing changes in C_n over trials. The following assumption seems reasonable: If response A_k occurs on a trial and is reinforced, then with probability μ the function v_k takes on the value one and that response moves to the top of the observing sequence. If the response is not reinforced, then with probability δ the v_k function for that response becomes zero and the observing sequence is reordered.

Given these assumptions it can be shown that for large n this model and the weak-strong model are equivalent. For example, in the two-response noncontingent case, if we let

(37)

$$S_1 = \ll 12 > , < 10 \gg \qquad S_2 = \ll 21 > , < 01 \gg$$

$$W_1 = \ll 21 > , < 00 \gg \qquad W_2 = \ll 21 > , < 01 \gg ,$$

then the transitions among states is that given by Eq. 2 and at asymptote the predictions for the weak-strong model are precisely those of the model outlined in this section.

The implications of the approach that we have just considered are broader than the fact that we achieve results identical to those derivable from weak-strong axioms. A number of models may be generated, starting with the notions of approach tendencies and observing vectors, if one examines various natural modifications of the conditioning and responding assumptions that were sketched above. In view of the possibility that some of these models will provide further insights into choice behavior, this frame of reference merits further investigation.

Appendix A

Listed below are the expressions for the asymptotic joint probabilities of the form $P(A_{1,n}E_{j,n-1}A_{k,n-1}E_{\ell,n-2}A_{m,n-2})$ for the weak-strong model $(j,k,\ell,m = 1,2)$. The conditional statistics may be obtained by noting that in the noncontingent case.

$$P(A_i|E_jA_kE_\ell A_m) = \frac{P(A_iE_jA_kE_\ell A_m)}{Pr(E_j)Pr(E_\ell)P(A_k|E_\ell A_m)P(A_m)}$$

where $P(E_1) = \pi$ and $P(E_2) = (1-\pi)$.

$$P(A_1E_1A_1E_1A_1) = \frac{\pi^2}{N^2}\{A+3(n-1)A^2+(N-1)(N-2)A^3\}$$

$$P(A_1E_1A_1E_1A_2) = \frac{\pi^2}{N^2}[u_3\delta+(N-1)A(B+2u_3\delta)+(N-1)(N-2)A^2B]$$

$$P(A_1E_1A_1E_2A_1) = \frac{\pi(1-\pi)}{N^2}[C+(N-1)A(A+2C)+(N-1)(N-2)A^3]$$

$$P(A_1E_1A_1E_2A_2) = \frac{\pi(1-\pi)}{N^2}\{(N-1)AB[1+(N-2)A]\}$$

$$P(A_1E_1A_2E_1A_1) = \frac{\pi^2}{N^2}\{(N-1)A[u_3\delta+B+(N-2)AB]\}$$

$$P(A_1E_1A_2E_1A_2) = \frac{\pi^2}{N^2}\{u_3\delta(1-\delta)+u_4\delta^2+(N-1)[2Bu,\delta+AD]+(N-1)(N-2)AB^2\}$$

$$P(A_1E_1A_2E_2A_1) = \frac{\pi(1-\pi)}{N^2}\{u_2\delta^2+(N-1)[A\delta(u_2+u_3)+BC]+(N-1)(N-2)A^2B\}$$

$$P(A_1E_1A_2E_2A_2) = \frac{\pi(1-\pi)}{N^2}[u_3\delta(1-\mu_3)+(N-1)B(u_3\delta+A)+(N-1)(N-2)AB^2]$$

$$P(A_1E_2A_1E_1A_1) = \frac{\pi(1-\pi)}{N^2}\{u_1+u_2[1-\delta(1-\mu)]+(N-1)A(2A+C)+(N-1)(N-2)A^3\}$$

$$P(A_1E_2A_1E_1A_2) = \frac{\pi(1-\pi)}{N^2}[u_3\delta(1-\delta)+(N-1)(BC+2Au_3\delta)+(N-1)(N-2)A^2B]$$

$$P(A_1E_2A_1E_2A_1) = \frac{(1-\pi)^2}{N^2}[u_1(1-\delta^2)+u_2(1-\delta)^2+3(N-1)AC+(N-1)(N-2)A^3]$$

$$P(A_1E_2A_1E_2A_2) = \frac{(1-\pi)^2}{N^2}(N-1)B[C+(N-2)A^2]$$

$$P(A_1E_2A_2E_1A_1) = \frac{\pi(1-\pi)}{N^2}(N-1)AB[1+(N-2)A]$$

$$P(A_1E_2A_2E_1A_2) = \frac{\pi(1-\pi)}{N^2}(N-1)[Bu_3\delta+AD+(N-2)AB]$$

$$P(A_1E_2A_2E_2A_1) = \frac{(1-\pi)^2}{N^2}(N-1)[BC+Au_2\delta+(N-2)A^2B]$$

$$P(A_1E_2A_2E_2A_2) = \frac{(1-\pi)^2}{N^2}(N-1)AB[1+(N-2)B]$$

where

$$A = u_1 + u_2$$

$$B = u_3 + u_4$$

$$C = u_1 + u_2(1-\delta)$$

$$D = u_3(1-\delta) + u_4$$

Appendix B

The values of n_{jk} presented in Table 10 are the total numbers of asymptotic trials on which E_j and A_k both occurred, pooling over all subjects in each group. These values are the denominators for the first-order conditional statistics presented in the paper. Values of $n_{jk\ell m}$ in Table 11 are the numbers of pairs of asymptotic trials containing E_j and A_k on trial n and E_ℓ and A_m on trial $n-1$. These values are the denominators of the second-order conditional statistics analyzed for the Friedman et al study.

Table 10

Values of n_{jk} for several studies

Experiment	Group	n_{11}	n_{21}	n_{12}	n_{22}
Suppes and Atkinson	Z	1238	595	602	365
	F	900	656	537	307
	T	1008	673	428	291
Myers et al	.6-0¢	590	395	358	241
	.7-0¢	1042	449	356	133
	.8-0¢	1382	340	197	61
	.6-1¢	771	516	413	280
	.7-1¢	1204	517	194	65
	.8-1¢	1463	368	115	34
	.6-10¢	828	582	349	221
	.7-10¢	1202	515	193	70
	.8-10¢	1489	393	85	13
Friedman et al		4815	1166	1028	268

Table 11

Values of $n_{jk\ell m}$ for the Friedman et al Study

$jk\ell m$	$n_{jk\ell m}$	$jk\ell m$	$n_{jk\ell m}$
1111	3435	2211	868
1112	585	2112	126
1121	699	2121	168
1122	82	2122	27
1211	427	2211	85
1212	259	2212	58
1221	225	2221	74
1222	109	2222	50

References

Anderson, N. H. Comments on Professor Estes' paper. In A. W. Melton (Ed.), Categories of human learning. New York: Academic Press, 1963.

Anderson, T. W. and Goodman, L. A. Statistical inference about Markov chains. Ann. Math. Stat., 1957, 28, 89-110.

Atkinson, R. C. The observing response in discrimination learning. J. exp. Psychol., 1961, 62, 253-262.

Atkinson, R. C. Choice behavior and monetary payoff: strong and weak conditioning. In J. H. Criswell, H. Solomon, and P. Suppes (Eds.) Mathematical methods in small group processes, Stanford: Stanford Univ. Press. 1962.

Atkinson, R. C. and Estes, W. K. Stimulus sampling theory. In R. D. Luce, R. R. Bush, and E. Galanter. (Eds.), Handbook of mathematical psychology. Vol. 2, New York: Wiley, 1963.

Bower, G. H. Choice-point behavior. In R. R. Bush and W. K. Estes (Eds.), Studies in mathematical learning theory. Stanford, California: Stanford University Press, 1959, 109-124.

Bush, R. R. and Mosteller, F. Stochastic models for learning. New York: Wiley, 1955.

Calfee, R. Extended training under probabilistic reinforcement schedules with rats. Univ. of Calif., Los Angeles Ph.D. thesis, 1963.

Cole, M. Search behavior: a correction procedure for three-choice probability learning. Indiana University, Ph.D. thesis, 1962.

Cotton, J. W. and Rechschaffen, A. Replication report: two and three choice verbal conditioning phenomena, J. exp. Psychol., 1958, 56, 96.

Edwards, W. Reward probability, amount, and information as determiners of sequential two-alternative decisions. J. exp. Psychol., 1956, 52, 177-188.

Estes, W. K. Toward a statistical theory of learning. Psychol. Rev., 1950, 57, 94-107.

Estes, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.), Studies in mathematical learning theory. Stanford, California: Stanford Univ. Press, 1959.

Estes, W. K. A random-walk model for choice behavior. In K. J. Arrow, S. Karlin, and P. Suppes (Eds.), Mathematical methods in the social sciences. Stanford, California: Stanford Univ. Press, 1960.

Estes, W. K. Theoretical treatments of differential reward in multiple-choice learning and two-person interactions. In J. H. Criswell, H. Solomon, and P. Suppes (Eds.), Mathematical models in small group processes, Stanford: Stanford Univ. Press, 1962.

Friedman, M. P., Burke, C. J., Cole, M, Estes, W.K., Keller, L. and Millward, R. B. Two-choice behavior under extended training with shifting probabilities of reinforcement. In R. C. Atkinson (Ed.), Studies in mathematical psychology. Stanford: Stanford Univ. Press, 1963.

Gardner, R. A. Probability-learning with two and three choices. Amer. J. Psychol., 1957, 70, 174-185.

Katz, L. Monetary incentive and range of payoffs as determiners of risk-taking. J. exp. Psychol., 1962, 64, 541-544.

LaBerge, D. A model with neutral elements. In R. R. Bush and W. K. Estes (Eds.), Studies in mathematical learning theory, Stanford: Stanford Univ. Press, 1959.

Luce, R. D. Individual choice behavior, New York: Wiley, 1960.

Myers, J. L. and Fort, J. G. A sequential analysis of gambling behavior. Paper presented at meeting of Psychonomic Society, New York, 1962.

Myers, J. L., Fort, J. G., Katz, L. and Suydam, M. M. Supplementary report: differential monetary gains and losses and event probabilities in a two-choice situation. J. exp. Psychol., 1963, in press.

Myers, J. L. and Ktaz, L. Range of payoffs and feedback in risk taking. Psychol. Rep., 1962, 10, 483-486.

Myers, J. L. and Sadler, E. Effects of range of payoffs as a variable in risk taking. J. exp. Psychol., 1960, 60, 306-309.

Savage, L. J. The foundations of statistics, New York: Wiley, 1954.

Siegel, S. Theoretical models of choice and strategy behavior - stable state behavior in the two-choice uncertain outcome situation. Psychometrika, 1959, 24, 203-216.

Suppes, P. and Atkinson, R. C. Markov learning models for multiperson interations. Stanford: Stanford Univ. Press, 1960.

Footnotes

1. Support for this research was provided by NIMH grant M-3803 and NSF grant G-22118 to the first author and by NIMH grant M-5184 to the second author. Computer time was inpart paid for by NIMH grant M-6154. Additional support was provided by a special fellowship awarded by the U. S. Public Health Service to the first author.

2. Values of n_{jk} are tabled in appendix B for all experimental groups discussed in this section. Thus the data on which our analyses have been based can be completely reproduced, and the interested reader may use the data to analyze alternative models.

3. If the investigator is only interested in predicting $P(A_1)$, estimates of φ can be obtained by direct solution of Eq. 7. The procedure, and the resulting fit (which is better than that reported in Table 3) are reported by Myers et al (1963).

4. It is possible that in the typical experimental situation the subject decides on his response prior to the signal to respond. Under these conditions response time, measured from the onset of the signal, would reflect the speed of reaction to the trial signal, and not choice time. A more sensitive test of response time predictions might be made if subjects were permitted to pace themselves; latency would be measured from the onset of the event on trial n to the occurrence of the response on trial $n + 1$.