6

Choice-Point Behavior[1]

GORDON H. BOWER, Stanford University

This paper presents a detailed analysis of the behavior of a subject at a choice point having several discriminative stimuli to which he may respond selectively. It has frequently been observed that a subject in such a situation spends some time sampling, or orienting to, the different stimuli before making his choice response. For example, in a black-white discrimination, a rat often looks back and forth repeatedly between the black stimulus and the white stimulus, as if to "make up his mind," before he approaches one stimulus or the other. Historically, psychologists have referred to orientation back and forth between discriminative stimuli as "vicarious trial and error" or VTE behavior.

It seems likely that our understanding of choice behavior and the variables affecting it will be enhanced by a detailed analysis of the component behaviors which constitute a choice response. Empirically, it seems that choice behavior in the situation described above may be analyzed into two component behaviors: orientation to the separate discriminative stimuli, and approach to these stimuli following orientation. Choices may be regarded as the outcome of the interaction of these underlying component responses. The models of this paper represent ways by which the approach and orientation tendencies to the several discriminative stimuli might combine to determine choices. In addition, the models provide a complete account of the sequences of orienting responses the subjects perform before approaching one stimulus or the other.

The models may be applied to those choice situations in which it is possible to identify orienting responses which serve separately to expose the subject to either of two discriminative stimuli. One simple example is a single-unit T maze with one black and one white arm. On every trial of the experiment, the black and white stimuli are separated spatially and the subject cannot orient to both simultaneously. To give concreteness to the assumptions of the model, the following discussion is referred to the single-unit T maze.

General Assumptions

The stimuli at the choice point of a T maze will be classified as follows: (1) the stimuli the subject samples when he is oriented straight ahead, which

[1] See footnote 1, p. 76.

we will denote by the set S_0; and (2) the stimuli, S_1 and S_2, available for sampling on the right or left side of the choice point when the subject orients in that direction. For example, in a black-white discrimination, S_1 may denote the stimuli of the black alley and S_2 the stimuli of the white alley. If the positions of the black-white cues are randomly interchanged over trials, then S_1 appears equally often on the left and right sides of the T maze.

The responses elicited by these stimuli will be described as follows: (1) S_0 elicits orienting responses to the left or right with respect to the subject's body position; (2) when the subject is oriented to S_1 he either approaches S_1 or reorients; when the subject is oriented to S_2 he either approaches S_2 or reorients.

From the sequences of orienting responses that occur on individual trials of an experiment, it is possible to estimate the probabilities of these responses in the presence of the appropriate stimuli. We let F_0 denote the probability that the subject orients to the right after he is oriented to S_0 and $1 - F_0$ the probability that S_0 elicits an orientation to the left. By F_1 we shall denote the probability of an approach response when the subject is oriented to S_1. Similarly, by F_2 we shall denote the probability of an approach response when the subject is oriented to S_2. For convenience, we shall refer to F_0, F_1, and F_2 as the *basic* probabilities.

Over training trials within an experiment, the basic probabilities will undergo changes due to learning. These trial-to-trial changes are not of major interest in the first section of this paper. Instead, the major interest is in testing the implications of the model for orienting behavior on a single experimental trial, given the values of F_0, F_1, and F_2.

The Markovian Assumptions

The assumptions about the choice-point stimuli and the responses elicited by these stimuli permit a Markovian representation of choice-point behavior. The states of the Markov process represent the significant positions of orientation the subject may take at the choice point. We let states 0, 1, and 2 represent the subject's orientation to S_0, S_1, and S_2, respectively. States 3 and 4 will represent the subject's final approach to S_1 or S_2, respectively. States 3 and 4 are "absorbing" states in the sense that choice-point behavior is terminated on a particular experimental trial when the subject approaches S_1 or S_2.

The word "transition" will be used to denote the movement of the subject from one state to another. An experimental trial starts with the subject in state 0 and ends when he enters state 3 or state 4. Between these events, the subject will perform some ordered sequence of orientations and reorientations, corresponding to transitions between states.

The basic probabilities represent transition probabilities among the several states of the system. Thus F_0 is the probability of a transition from state 0 to state 1 if S_1 is on the right side of the maze on that trial, or from state 0 to state 2 if S_2 is on the right on that trial; F_1 is the probability of a transition from state 1 to state 3; and F_2 is the probability of a transition from state 2 to state 4.

In representing choice-point behavior as a Markov process, two assumptions are made: (1) the transition behavior is only one-step dependent—that is, the probability of a transition from state i to state j is independent of the path by which the system entered state i; and (2) the transition probabilities F_0, F_1, and F_2 remain constant for the duration of a single experimental trial.

With these assumptions, we will now consider one particular Markov representation of choice-point behavior. A few implications of this representation will be compared with some experimental results.

Model A

In this model central importance is assigned to state 0. It is assumed that if the subject is oriented to S_1 or S_2 and then turns away from this stimulus, he returns to state 0 (orients to S_0), and S_0 elicits a second orientation to the left or right. A diagram of this process is shown in Fig. 1 for a trial on which S_1 and S_2 are on the right and left side of the maze, respectively.

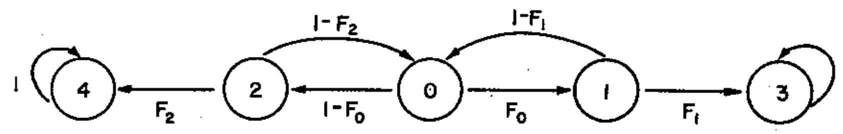

FIGURE 1. The Markovian representation of model A.

Pointed arrows represent the possible one-step transitions. The probabilities of the transitions are indicated beside the arrowed lines. This diagram is equivalent to a 5×5 matrix of transition probabilities. States 3 and 4 are absorbing states; the probability is 1 that the subject remains in these states once he enters them on any trial. The definition of a trial in a T maze ensures that the subject starts every trial at state 0.

Every experimental trial is eventually terminated when the subject approaches S_1 or S_2. To derive the probability that the subject approaches S_1 (is absorbed at state 3), we note that from state 0 to state 3 there is a direct path through state 1. The probability that the subject follows this path is F_0F_1. However, the subject may perform "loops" of orienting responses of the form 0, 1, 0, or 0, 2, 0 before he takes the final path 0, 1, 3. The probability of a 0, 1, 0 loop is $F_0(1 - F_1)$, and the probability of a 0, 2, 0 loop is $(1 - F_0)(1 - F_2)$. The subject may perform an indefinite number of these loops before he takes the final path 0, 1, 3. Therefore, the probability of absorption at state 3 may be written as the infinite sum

$$\text{(1)} \qquad \Pr(3) = F_0F_1 \sum_{i=0}^{\infty} [F_0(1 - F_1) + (1 - F_0)(1 - F_2)]^i = \frac{F_0F_1}{F_0F_1 + (1 - F_0)F_2}.$$

This last expression gives the probability that a subject starting in state 0 is eventually absorbed at state 3 (approaches S_1). The probability that the subject approaches S_2 is $\Pr(4) = 1 - \Pr(3)$.

Equation 1 does not constitute a test of the model if estimates of the basic probabilities are first obtained from the data. All the information in the data relevant to $\Pr(3)$ is exhausted in obtaining the estimates of the basic probabilities.

To show this from probability considerations alone, let A and $\bar{A}$ denote the events of orienting to the right or left, respectively. Events B and C will denote approach to right or left, respectively. Without recourse to the model, we may write the analytic expression for the proportion of approaches to the right as

$$(2) \qquad \Pr(3) = \frac{p(A)p(B \mid A)}{p(A)p(B \mid A) + p(\bar{A})p(C \mid \bar{A})} .$$

In the model, the estimates of the basic probabilities are

$$(3) \qquad F_0 \triangleq p(A), \quad F_1 \triangleq p(B \mid A), \quad F_2 \triangleq p(C \mid \bar{A}) .$$

When these estimates are substituted into Equation 1, Equation 2 follows identically; therefore, all information relevant to $\Pr(3)$ has been used in obtaining the estimates of F_0, F_1, and F_2.

This result cautions us to check every statistic derived from the model to determine whether all information relevant to the statistic has been exhausted by the estimates of F_0, F_1, and F_2. If these estimates do not exhaust all information relevant to the statistic, then the statistic may be considered as a prediction to use in testing the model.

The model may be tested by some of its predictions about the sequences of orienting responses. On any given trial, a subject will perform some particular sequence of orienting responses. A sample protocol of a subject's behavior might be "0, 1, 0, 1, 0, 2, 4." One way to study these sequences is to count the number of times certain patterns of transitions occur. For example, the pattern 0, 1, 0 occurs twice in the protocol above. If the model is correct, it should predict the distribution of the number of occurrences of any pattern of transitions.

As one example, consider the "loop" patterns 0, 1, 0 and 0, 2, 0. Let L, for loops, be a random variable denoting the number of loops that occur on a given trial for one subject. The variable L may have any of the values 0, 1, 2, $\cdots$ of the nonnegative integers. The problem is to derive from the model the distribution of L, i.e., the probability that L takes on some arbitrary value k.

The first few terms of this distribution may be written down by inspection of Fig. 1. The probability that $L = 0$ is given by the probability that the subject is absorbed at 3 or 4 in two steps, which is $F_0F_1 + (1 - F_0)F_2$. The probability that $L = 1$ is given by the probability that the subject performs exactly one loop and then is absorbed in two steps. The probability of one loop or the other is $F_0(1 - F_1) + (1 - F_0)(1 - F_2)$, and the probability of absorption in two steps following this loop is $F_0F_1 + (1 - F_0)F_2$. The probability that L is exactly 1 is given by the product of these two probabilities. After writing a few more terms for $L = 2$, $L = 3$, $\cdots$, it becomes apparent that the series follows the geometric law given by

$$(4) \qquad \Pr(L = k) = [F_0F_1 + (1 - F_0)F_2][F_0(1 - F_1) + (1 - F_0)(1 - F_2)]^k .$$

Equation 4 gives the probability distribution of the number of loops before absorption occurs. If we set $J = F_0F_1 + (1 - F_0)F_2$, we may write Equation 4 more simply as

$$(5) \qquad \Pr(L = k) = J(1 - J)^k .$$

Equation 5 gives $\Pr(L = k)$ as an exponentially decreasing function of k and is the well-known geometric distribution. In testing the model the empirical distribution of L will be compared with the distribution predicted by the model.

As a second property of orientation sequences, suppose we count the total number of transitions in both directions between states 1 and 2. Let V denote the number of transitions between states 1 and 2. The derivation of the distribution of V will not be presented, since it is rather long. The result which will be used in testing the model is the average, or expected, value of V. The expected value of V derived from the model is given by

$$(6) \qquad E(V) = \frac{F_0(1 - F_0)[(1 - F_1) + (1 - F_2)]}{F_0F_1 + (1 - F_0)F_2} .$$

This expression provides a second prediction from the model to compare with experimental results.

Some Experimental Results

In testing some hypotheses concerning free versus forced choice trials, Milton A. Trapold and the writer recorded the VTE behavior of rats in a place + response learning situation. The experiment employed fifteen pairs of hungry rats run in an all-white T maze with five Noyes food pellets (about .25 gm.) on the right side and one pellet of food on the left side. Prior to the T-maze training, each subject was placed several times in both goalboxes and allowed to eat one pellet of food. Goalbox training was continued until each subject ate promptly upon being placed in either goalbox. Then four spaced trials a day were given for eleven days, after which all subjects were regularly choosing the five-pellet side of the maze.

Each pair of animals consisted of an experimental subject and his "yoked" control mate. On trials 1, 2, and 4 of each training day, the experimental subject of the pair was given free choices (noncorrection) while his control mate was forced to the same side the experimental subject chose on those trials. Forcing was accomplished by lowering a guillotine door at the choice point, blocking one of the maze arms. On the third trial of the day, the control subject was given a free choice and his experimental mate was forced to the side the control subject chose. This procedure ensured that following every trial the experimental subject and his yoked control mate were matched on the sequence of left and right turns and reinforcing events.

A mirror placed three feet above the choice point permitted observation of choice-point behavior. On each free-choice trial, the choice-point behavior of each rat was recorded in terms of his sequence of orienting responses. A typical protocol might read "0, 2, 0, 1, 3," describing a subject that entered the choice point, looked to the left, to the center, and to the right, and then approached the right arm of the maze. The orienting responses were discrete and clearly observable. The two experimenters agreed perfectly in recording the sequences of orienting responses.

The forced-choice versus free-choice comparison will not concern us here. Briefly, it was found that (1) forced trials definitely had a learning effect which appeared in measures of the basic probabilities, and (2) in comparison with a free trial, a forced trial in this situation had a slightly smaller effect on these measures than did an equivalent free trial.

The data of primary interest here are the sequences of orienting responses. After estimates of F_0, F_1, and F_2 are obtained, the predictions of the model about these orientation sequences will be tested.

TABLE 1

Observed Values of F_0, F_1, and F_2 (Experimental Group Only) Blocks of Three Free Trials

Trial Block	F_0	F_1	F_2
1	.50	.55	.37
2	.55	.71	.35
3	.59	.78	.32
4	.69	.93	.26
5	.78	.93	.23
6	.90	.98	.20
7	.90	.98	.20
8	1.00	1.00	—*
9	.98	1.00	0
10	.92	.98	0
11	.98	1.00	0

* No subject looked left during this block of trials. Consequently, F_2 could not be estimated.

In Table 1 the estimated values of F_0, F_1, and F_2 for the experimental group (fifteen subjects) are given in blocks of three free trials. These estimates are easily obtained from the data. For example, F_1 for the nth block of trials is obtained by dividing the number of right turns that occurred during that trial block by the total number of times the subjects looked to the right during that trial block. The estimate of F_0 is given by the proportion of times the subjects looked to the right after they had been oriented straight ahead. By substituting the values of F_0, F_1, and F_2 on the

nth trial block into Equation 1, the probability of a right turn for that trial block is obtained.

In Table 1 both F_0 and F_1 appear to have asymptotes of unity, while F_2 has an asymptote of zero. The corresponding functions for the control group (not shown) have the same asymptotes and approximately the same shapes. The last four values for F_2 are based on small numbers of observations because the subjects made fewer and fewer orientations to the small-reward side.

As the first test of the model, we consider the distribution of L, the number of loops of the form 0, 1, 0 or 0, 2, 0. The model implies that L should have the geometric distribution given by Equation 5, i.e.,

$$\Pr(L = k) = J(1 - J)^k .$$

The results of the experiment show that the daily distributions of L do have a geometric form. The sequence of daily geometric distributions tends to a point distribution with all density at $L = 0$ (corresponding to F_0 and F_1 both reaching asymptotes of unity in Equation 4).

There is one marked deviation from the predicted form of the distribution of L. The distribution on day 1 of the experiment departs significantly from the geometric form for both groups ($P < .01$, Kolmogorov-Smirnov one-sample test). On day 1, the relative frequency of $L = 1$ was larger than the relative frequency of $L = 0$, although all other values were of the expected order of magnitude. This result suggests that the model may be wrong in its description of behavior during the initial trials of the experiment.

If day 1 is excluded and the observations of L are pooled over days 2 through 7 to obtain a large sample, the geometric form of the distribution is clear. In Table 2, the empirical and predicted distributions are shown for the experimental and control groups, respectively. The constant J was estimated by the first point on the empirical distribution (setting $k = 0$ in Equation 5), and the remaining values of the theoretical distribution were calculated from Equation 5 with this estimate substituted for J. In both cases, the geometric distribution fits the data very well.

TABLE 2

Distribution of L

L	Experimental Group		Control Group	
	Observed	Predicted	Observed	Predicted
0	.710	.710	.733	.733
1	.226	.206	.200	.196
2	.041	.060	.059	.052
3	.015	.017	.008	.014
4	.011	.005	0	.004
5	0	.001	0	.001

Further information on the goodness-of-fit may be obtained by comparing the first few moments of the observed and predicted distributions. In Table 3 the first three raw moments are given for the experimental and control groups, respectively. Over all, the predicted values are quite close to the observed values.

TABLE 3

Moments of L Distributions

Moment	Experimental Group		Control Group	
	Observed	Predicted	Observed	Predicted
First (Mean)	.40	.41	.34	.34
Second	.70	.75	.51	.59
Third	1.65	1.30	.89	.98

Note.—The mth raw moment is given by $\sum_{k=0}^{\infty} k^m \Pr(L = k)$.

A second test of the model is the prediction of the average number of transitions between states 1 and 2. Predictions of $E(V)$ over trials were obtained by substituting the successive values of F_0, F_1, and F_2 from Table 1 into Equation 6. In Table 4 the obtained and predicted values of $E(V)$ are compared over trials.

TABLE 4

Trial Changes in Observed and Predicted Values of $E(V)$ by Blocks of Three Free Trials (Experimental Group Only)

Trial Block	$E(V)$	
	Observed	Predicted
1	1.13	.59
2	.76	.42
3	.62	.37
4	.38	.24
5	.24	.19
6	.09	.08
7	.11	.08
8	0	0
9	.02	.02
10	.11	.08
11	.02	.02

It is clear from Table 4 that the predictions are considerably below the observed values of $E(V)$. The predictions for the first four blocks of trials are roughly half as large as the observed values.

The source of this discrepancy is clear. If the subjects were oriented to S_1 and turned away, then they were very likely to orient to S_2 rather than to look back to S_1 again. This observation is supported by further analysis

of the data. The orientation sequences of the first four blocks of trials were analyzed for repetition patterns (1, 0, 1 or 2, 0, 2) and alternation patterns (1, 0, 2 or 2, 0, 1). The frequencies of the alternation patterns for the four blocks of trials were 51, 34, 28, and 17, respectively, while the corresponding frequencies of the repetition patterns were 2, 3, 3, and 0. The subjects clearly show a higher degree of alternation of orientations to S_1 and S_2 than is predicted by the model. This indicates that transitions from state 0 do not conform to the path-independence condition. The next orientation of the subject from state 0 depends markedly on whether he was oriented to the left or right before coming back to the center (state 0).

In the next section, we shall consider a revised Markovian representation which does not encounter the "alternation" difficulty of model A. In model B the states 0, 1, and 2 are redefined in such a way that state 0 plays a minor role in the model. State 0 is redefined to be the position of the subject before he makes his first orientation to S_1 or S_2 on a given trial. State 1 is redefined so that head movements from right to center are ignored. The subject is said to leave state 1 only when he approaches state 3 or reorients completely to the other side (state 2). State 2 is defined in a similar manner.

One example of the sequences to be analyzed by model B is 0, 1, 2, 1, 2, 4. Such sequences may be obtained from the earlier protocols by deleting all 0's beyond the first and collapsing all runs of 1's or 2's separated only by zeros. Thus, the earlier protocol 0, 1, 0, 1, 0, 2, 4 will be rewritten as 0, 1, 2, 4. This revised sequence is then ready for analysis by model B.

Model B

The Markovian representation of model B is diagramed in Fig. 2. The new assumption of this model is that the subject orients back and forth between S_1 and S_2 and uses S_0 only to determine the starting point of the cycles of reorientation.

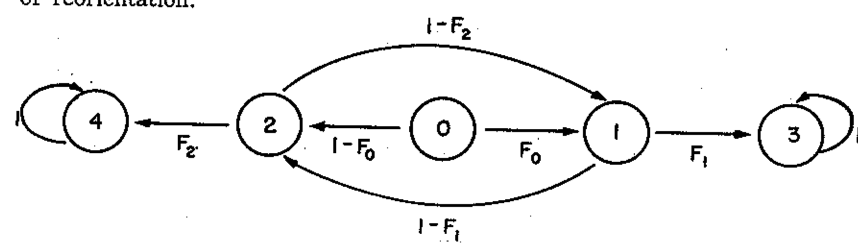

FIGURE 2. The Markovian representation of model B.

The estimates of F_0, F_1, and F_2 for model B differ slightly from the estimates in model A: F_0 is estimated by the proportion of times that subjects orient to the right upon first reaching the choice point; F_1 and F_2 are estimated by the proportion of approach responses per orientation in the S_1 or S_2 direction, respectively.

Four results will be derived from the model: (1) the unconditional probability, Pr (3), that the subject is absorbed at 3; (2) the conditional probabilities, Pr (3 | 1) and Pr (3 | 2), that subjects are eventually absorbed at 3 given that their first orientation is to the right or left, respectively; (3) the probability distribution of V, the number of transitions between states 1 and 2; and (4) the average value of V, $E(V)$. The results for Pr (3) and $E(V)$ cannot be used to test the model since the estimates of F_0, F_1, and F_2 in model B exhaust all information relevant to both statistics. They are derived here for use in a later section of the paper. The model will be tested by its predictions about Pr (3 | 1), Pr (3 | 2), and the distribution of V.

Consider first the conditional probability that the subject is eventually absorbed in State 3 given that he orients first to the right (to S_1, with the arrangement in Fig. 2). From state 1 the subject is absorbed at 3 in the next step with the probability F_1. However, the subject may make an indefinite number of loops of the form 1, 2, 1 before he finally takes the step from state 1 to state 3. The probability of exactly j loops of the form 1, 2, 1 starting from state 1 is $F_1[(1-F_1)(1-F_2)]^j$. From these considerations, we may write Pr (3 | 1) as the infinite sum

$$(7)\qquad \Pr(3 \mid 1) = F_1 \sum_{j=0}^{\infty} [(1-F_1)(1-F_2)]^j = \frac{F_1}{F_1+F_2-F_1F_2}.$$

Equation 7 gives an expression for the probability of eventual absorption at 3, conditional on the subject's looking toward S_1 first upon reaching the choice point.

If the subject starts the trial by orienting to S_2, then with probability $1-F_2$ a transition to state 1 occurs. Once the subject enters state 1, the probability of eventual absorption at 3 is given by Equation 7. Hence, we may write

$$(8)\qquad \Pr(3 \mid 2) = (1-F_2)\Pr(3 \mid 1) = \frac{(1-F_2)F_1}{F_1+F_2-F_1F_2}.$$

Using Equations 7 and 8 we may write the unconditional probability of absorption at state 3 as

$$(9)\qquad \Pr(3) = F_0 \Pr(3 \mid 1) + (1-F_0)\Pr(3 \mid 2) = \frac{F_1[F_0+(1-F_0)(1-F_2)]}{F_1+F_2-F_1F_2}.$$

The model may be tested by Equations 7 and 8. To estimate Pr (3 | 1) on each trial block, we select all subjects that looked to the right first upon entering the choice point and calculate the proportion of these subjects that were eventually absorbed at 3. A similar estimate of Pr (3 | 2) is readily obtained from the data. The observed and predicted values of Pr (3 | 1) and Pr (3 | 2) are compared in Table 5. The standard deviations of the observed estimates are included to facilitate comparison. In every case, the predicted value is reasonably close to the observed value.

Next, we consider the distribution of V, the number of transitions between states 1 and 2 before absorption occurs. The expression for the probability

TABLE 5

Observed and Predicted Values of Pr (3/2) and Pr (3/1)
(Experimental Group Only)

Trial Block	Pr (3/2)			Pr (3/1)		
	Observed	Predicted	S.D.	Observed	Predicted	S.D.
1	.52	.47	.10	.70	.76	.10
2	.64	.57	.10	.80	.89	.09
3	.59	.62	.10	.96	.93	.04
4	.81	.72	.10	.97	.99	.03
5	.80	.74	.13	1.00	1.00	0
6	.80	.80	.18	1.00	1.00	0
7	.75	.80	.22	1.00	1.00	0
8	1.00	1.00	0	1.00	1.00	0
9	1.00	1.00	0	1.00	1.00	0
10	1.00	1.00	0	1.00	1.00	0
11	1.00	1.00	0	1.00	1.00	0

that V has some value k will depend on whether k is an even or odd integer. The subject can make k transitions from 1 to 2 in either direction by circling the "loop" 1, 2, 1 or 2, 1, 2 on $k/2$ occasions. The probability of one of these loops is $(1-F_1)(1-F_2)$. The probability of $k/2$ loops is this factor raised to the power $k/2$. If k is even, the loops begin and end in state 1 with probability F_0F_1, and begin and end in state 2 with probability $(1-F_0)F_2$. When k is odd, the subject is absorbed on the side opposite his initial orientation after making $(k-1)/2$ loops. From these considerations we may write the distribution of V as

$$(10)\qquad \Pr(V=k) = \begin{cases} [F_0F_1+(1-F_0)F_2][(1-F_1)(1-F_2)]^{k/2} & \text{if } k \text{ is even} \\ [F_0(1-F_1)F_2+(1-F_0)(1-F_2)F_1][(1-F_1)(1-F_2)]^{(k-1)/2} & \text{if } k \text{ is odd.} \end{cases}$$

The predicted distribution of V was compared with the obtained daily distributions of V. With the exception of day 1, the daily distributions conformed closely to the predicted form, tending over training to a point distribution with all density at $V = 0$. On day 1, the obtained and predicted distributions differed significantly ($P < .01$, Kolmogorov-Smirnov one-sample

TABLE 6

Observed and Predicted Values of V from Model B
(Experimental Group Only)

V	Observed	Predicted
0	.715	.726
1	.225	.211
2	.041	.046
3	.015	.013
4	.004	.003
5	0	.001

test). The relative frequencies of $V = 0$ and $V = 1$ on day 1 were in the reverse order from their predicted magnitudes.

By excluding day 1 and pooling observations on V over days 2 through 7, the distribution in Table 6 is obtained. The predicted distribution is given for comparison. The values of F_0, F_1, and F_2 were estimated from the pooled data.

The final result to be derived in this section is the average value of V. This may be calculated from the distribution of V given in Equation 10. After simplification, the result of this calculation is

$$E(V) = \frac{[(1 - F_2)(2 - F_1)] - F_0(F_1 - F_2)}{F_1 + F_2 - F_1F_2} . \tag{11}$$

This result will be used in the next section to describe the trial-to-trial changes in $E(V)$ implied by certain conditioning assumptions.

To summarize the results of testing model B, it was found that it correctly predicts the distribution of V and the trial-to-trial changes in $\Pr(3|1)$ and $\Pr(3|2)$. The outstanding discrepancy is the difference between the predicted and obtained distributions of V on day 1. Whether this is a sampling error or whether it is the general rule that the model is wrong for the very early trials (e.g., because of failure to take into account visual "exploration" habits) can only be discovered through further experimentation.

Learning Assumptions

In the preceding sections, we have tested some implications of the models for performance at a choice point, using estimates of the basic probabilities obtained from the choice data. In this section, we propose some conditioning assumptions to describe the trial-to-trial changes in the basic probabilities as a function of reinforcement experience.

It should be emphasized that the performance models of the earlier sections and the conditioning assumptions of this section are independent. The following learning assumptions are derived from statistical learning theory [2]. Presumably, at this point the Hullian theorist would prefer to derive the basic probabilities from corresponding reaction potentials which change in a computable manner with reinforcement. Regardless of the rule used in deriving the changes in the basic probabilities during learning, the performance models are available for general use in testing any conditioning assumptions.

In the following, S_0, S_1, and S_2 are identified with sets of stimulus events, or elements, which may be effective when the subject orients to S_0, S_1, and S_2, respectively. We let θ_0 represent the sampling probability of S_0 elements given an orientation to S_0; and we let θ represent the sampling probabilities of S_1 and S_2 elements given an orientation to S_1 or S_2, respectively. By F_0 we denote the proportion of F_0 elements conditioned to orienting right, by F_1 the proportion of S_1 elements conditioned to the approach response, and by $1 - F_1$ the proportion of S_1 elements conditioned to the response of reorienting (to S_2 in model B). A similar interpretation is provided for F_2.

With this as background, the conditioning assumptions may be stated. Consider first the conditioning of S_0 elements to the response of orienting right. If, on trial n, the subject is reinforced for going to the right (which he does just after orienting in that direction), or is nonreinforced for going to the left, then $F_{0,n}$ increases according to

$$F_{0,n+1} = (1 - \theta_0)F_{0,n} + \theta_0 . \tag{12}$$

If, on trial n, the subject is reinforced for going to the left, or is nonreinforced for going to the right, then $F_{0,n}$ decreases according to

$$F_{0,n+1} = (1 - \theta_0)F_{0,n} . \tag{13}$$

To illustrate these conditioning assumptions, consider an experiment in which the subject is forced to the right and left sides of the T maze on random proportions ϕ and $1 - \phi$ of the trials, respectively. Further, let a_1 and a_2 be the probabilities that reinforcement follows a forced choice to the right or left side, respectively. Under these conditions, $\phi a_1 + (1 - \phi)(1 - a_2)$ is the probability of applying the operator specified by Equation 12, and $\phi(1 - a_1) + (1 - \phi)a_2$ is the probability of applying the operator specified by Equation 13. The average value of F_0 on trial $n + 1$ may be written as the average of Equations 12 and 13:

$$F_{0,n+1} = (1 - \theta_0)F_{0,n} + \theta_0[\phi a_1 + (1 - \phi)(1 - a_2)] . \tag{14}$$

The solution to this difference equation is given by

$$F_{0,n} = F_{0,\infty} - (F_{0,\infty} - F_{0,1})(1 - \theta_0)^{n-1} , \tag{15}$$

where $F_{0,\infty} = \phi a_1 + (1 - \phi)(1 - a_2)$, and $F_{0,1}$ is the subject's initial probability of orienting to the right in the presence of S_0 stimuli.

The conditioning assumptions for F_1 and F_2 are similar to those for F_0. If reinforcement occurs following an approach to S_i $(i = 1, 2)$, then F_i increases according to

$$F_{i,n+1} = (1 - \theta)F_{i,n} + \theta . \tag{16}$$

If nonreinforcement occurs following an approach to S_i, then F_i decreases according to

$$F_{i,n+1} = (1 - \theta)F_{i,n} . \tag{17}$$

Finally, if the subject does not approach S_i on trial n, then F_i remains unchanged. That is, we apply the identity operator if the subject has no experience of reinforcement or nonreinforcement at S_i on trial n. For example, if a forced trial to S_1 occurs on trial n, the value of F_2 is not affected by this event (i.e., $F_{2,n+1} = F_{2,n}$).

To illustrate these assumptions, consider again the forcing situation described above. Suppose that S_1 (black alley) is always on the right and S_2 (white alley) on the left. This arrangement will be called "place + response" contingencies. The parameter ϕ is the probability that the subject is forced to S_1, and a_1 is the probability of a reinforcement on a forced trial to S_1. The probability of applying to $F_{1,n}$ the operator in Equation 16 is ϕa_1, and the probability of applying the operator in Equation 17 is $\phi(1-a_1)$. The probability that $F_{1,n}$ remains unchanged is $1-\phi$ since this is the probability of a forced trial to S_2. From these considerations we may write the average F_1 on trial $n+1$ as

$$F_{1,n+1} = (1-\phi\theta)F_{1,n} + \phi\theta a_1 . \tag{18}$$

By similar reasoning, the average value of $F_{2,n+1}$ is found to be

$$F_{2,n+1} = [1-(1-\phi)\theta]F_{2,n} + \theta(1-\phi)a_2 . \tag{19}$$

Equations 18 and 19 are difference equations which have the solutions

$$F_{1,n} = a_1 - (a_1 - F_{1,1})(1-\phi\theta)^{n-1} , \tag{20}$$

$$F_{2,n} = a_2 - (a_2 - F_{2,1})[1-(1-\phi)\theta]^{n-1} . \tag{21}$$

Equations 20 and 21 show that the asymptotes of F_1 and F_2 are a_1 and a_2, the probabilities of reinforcement for approach to S_1 and S_2, respectively. The rate of convergence to these asymptotes is $\phi\theta$ for F_1 and $(1-\phi)\theta$ for F_2.

The expressions derived above for the changes in the basic probabilities in the forced-trial situation may be checked by free-choice test trials. From the results of a free-choice test trial, each of the basic probabilities may be estimated directly and compared with the predicted changes specified by Equations 15, 20, and 21.

The basic probabilities may also be substituted into Equations 9 and 11 of model B to predict the trial-to-trial changes in Pr (3) and $E(V)$. For example, suppose in the forcing situation described that $a_1 = a_2 = 1$ and $\phi = 2/3$. Equations 20 and 21 imply that F_1 and F_2 approach asymptotes of unity, and Equation 15 implies that F_0 approaches an asymptote of ϕ. By substitution in Equation 9, these values imply that Pr (3) approaches an asymptote of ϕ. If the initial conditions are $F_{0,1} = .50$ and $F_{1,1} = F_{2,1}$, we may derive the result that Pr (3) starts at .50, rises with training to a maximum above 2/3, and then falls to an asymptote of 2/3.

By substituting the expressions for $F_{0,n}$, $F_{1,n}$, and $F_{2,n}$ into Equation 11, one can derive the prediction that, if $a_1 = a_2 = 1$, then $E(V)$ is a decreasing function of training trials and has an asymptote of zero. It is easy to see the basis for the latter result. Since F_1 and F_2 approach 1, in the limit the subject is absorbed following his first orientation of each trial; hence, $E(V) = 0$.

In the design above, the black alley (S_1) is always on the right and the white alley (S_2) on the left; however, one may randomly interchange the black and white alleys in their left-right positions from trial to trial. In this case, the experimenter may correlate any of several cues with the side on which forced trials and reinforcements are given. In the one case, which we shall call "place" contingencies, the experimenter ignores the left-right dimension and defines ϕ, a_1, and a_2 relative to the black-white dimension. Then ϕ is the probability that the subject is forced to S_1, and a_1 is the probability of reinforcement when the subject is forced to S_1. A second case, called "response" contingencies, arises when the experimenter ignores the black-white dimension and defines ϕ, a_1, and a_2 relative to the right-left positional dimension. In this case, ϕ is the probability of forcing the subject to the right, and a_1 is the probability of reinforcement on a forced trial to the right.

The model may be applied to situations involving "place" and "response" contingencies. One must recall that F_0 is defined as the probability of orienting to whatever stimuli are on the right side of the maze. Thus, the probability of an orientation to S_1 is F_0 if S_1 is on the right, but it is $1-F_0$ if S_1 is on the left. The Markov expressions (e.g., Equations 9 and 11) in the first section were derived for trials with S_1 on the right. For trials with S_1 on the left, the expressions are correct if F_0 and $1-F_0$ are interchanged wherever they appear in the equations.

To illustrate briefly the way the model handles "place" contingencies, we consider the case in which the subject is forced to S_1 with probability ϕ, and S_1 is on the left side with probability .50. Here a_1 and a_2 are the probabilities of reinforcement when the subject is forced to S_1 and S_2, respectively. The conditioning rules imply that in such circumstances the asymptote of F_0 will be .50, while the asymptotes of F_1 and F_2 are a_1 and a_2, respectively. If $a_1 = a_2 = 1$ and the initial conditions are as before, Equation 9 implies that the probability of approach to S_1 starts at .50, deviates to a maximum with training, then returns to an asymptote of .50. Note that this asymptotic result differs from the analogous result with the place + response contingencies described above. The difference is due to the different asymptotes of F_0 under the two sets of contingencies. The asymptote is $\phi = 2/3$ in the case of place + response learning, and it is .50 in the present case of place contingencies.

As an additional illustration of the conditioning assumptions, we may consider "response" contingencies such that ϕ is the probability of a forced trial to the right, and .50 is the probability that S_1 is on the right on any given trial of the experiment. Here we identify a_1 and a_2 as the probabilities of reinforcement when the subject is forced to the right and left, respectively. Under these conditions, the asymptotes implied by the conditioning assumptions are

$$F_{0,\infty} = \phi a_1 + (1-\phi)(1-a_2) , \tag{22}$$

$$F_{1,\infty} = F_{2,\infty} = \phi a_1 + (1-\phi)a_2 . \tag{23}$$

If $a_1 = a_2 = 1$ and we use the same initial conditions as before, then the probability of a right turn starts at .50 and deviates gradually to a limiting value of ϕ, the asymptote of F_0. In this case, the choice probability does not temporarily "overshoot" the eventual asymptote as it did in the two preceding cases.

A novel prediction can be made for the "response" contingencies when $a_1 = 1$ and $a_2 = 0$. That is, the subject is always reinforced when forced to the right side of the maze, but is never reinforced when forced to the left. With $a_1 = 1$ and $a_2 = 0$, Equations 22 and 23 imply that F_0 approaches an asymptote of unity, while F_1 and F_2 approach asymptotes of ϕ. When these values are substituted into Equation 9, the predicted asymptotic probability of a right turn is

$$\Pr(\text{Right}) = \frac{1}{2-\phi}\,. \tag{24}$$

By substituting these same values into Equation 11, we see that the average value of V should decrease to an asymptote given by

$$E(V) = \frac{1-\phi}{\phi}\,. \tag{25}$$

These two predictions seem counter to common-sense expectations. Intuitively, we might expect the subject to learn to ignore the black-white cues [3] and learn simply to "turn right." If they should be verified, these predictions would provide convincing evidence for the conditioning assumptions.

A note is appropriate here to explain the requirement of a forcing procedure for the previous experimental designs. A forcing precedure allows us to specify the probabilities with which certain events (e.g., an approach to left which is nonreinforced) occur on any given trial. Using these probabilities, we can derive mean curves for $F_{0,n}$, $F_{1,n}$, and $F_{2,n}$. By contrast, a free-choice procedure implies that the sequences of events the subjects experience will depend on the choices they make. Because these choices cannot be specified in advance, the probabilities of certain events are unknown before the experiment is performed. In such a case, closed expressions for the mean values of the basic probabilities cannot be derived from the conditioning assumptions; however, they may be computed by Monte Carlo techniques [1]. In principle, then, the model may also be applied to experimental situations involving only free-choice trials.

The foregoing discussion has illustrated some implications of the conditioning assumptions. In principle, the assumptions are relevant to a large body of learning problems. In every case, the predictions of the trial-to-trial changes in choice probability depend jointly on the conditioning assumptions and the Markovian representation. Fortunately, these two classes of assumptions can be tested independently by a single set of experimental results if the experimenter records the sequence of orienting responses exhibited by the subject at the choice point. Thus the model may encourage experimenters to look more closely at these underlying components of choice behavior. It may be at this level of analysis that we will discover sufficient lawfulness to reward our investigations of choice behavior.

REFERENCES

1. Bush, R. R., and Mosteller, F. *Stochastic models for learning*. New York: Wiley, 1955.
2. Estes, W. K., and Burke, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–86.
3. Restle, F. Discrimination of cues in mazes: a resolution of the "place-vs-response" question. *Psychol. Rev.*, 1957, **64**, 217–28.