MATHEMATICAL MODELS IN RESEARCH ON PERCEPTION AND LEARNING

by

Richard C. Atkinson

TECHNICAL REPORT NO. 43

December 25, 1961

PSYCHOLOGY SERIES

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

Mathematical Models in Research on Perception and Learning[1,2]

Richard C. Atkinson

Stanford University

Introduction

The purpose of this paper is to examine the role of mathematical models in research on perception and learning. The reader, however, should be warned at the outset that we will not present a formal philosophical analysis of the function of models but, instead, will examine the development and application of a specific model. A formal analysis would lead us into a great deal of abstract discussion and would not stress current developments in research techniques. The model that will be examined deals with a forced-choice signal detection situation. It is particularly useful for illustrative purposes because it combines two quite distinct processes: a simple perceptual process and a learning process. As the theory is developed, we will be able to indicate the role of mathematical models in determining programs of psychological research and in specifying the types of empirical observations to be made.

---

[1]The ideas presented in this paper have been much influenced by discussions with R. Kinchla and E.C. Carterette. The research was supported by the National Institute of Mental Health under Contract M-5184. Portions of the paper were presented at an International Colloquium on the concept and the role of models in mathematics and science held in Utrecht on January 4-8, 1960.

[2]This paper was prepared as a contribution to a revised edition of Psychological Theory edited by Melvin H. Marx.

Before turning to the example, a few general comments seem in order. The use of mathematical models is virtually synonomous with the construction of a quantitative theory of behavior. From a mathematical standpoint it is logically possible to have a theory of behavior that leads only to qualitative predictions. However, in the history of science it is difficult to find theories of this sort that have had sustained empirical significance. From the systematic standpoint a theory based only on qualitative distinctions leads to a relatively small number of testable predictions. Further, as the set of phenomena that we study expands in complexity so also does the reasoning necessary for the design of experiments and the formulation of hypotheses. Ordinary logic becomes inadequate and the elaboration of the theory requires the powerful tool of mathematical analysis.

In this regard, perhaps the most important role of mathematical models in recent psychological research has been to provide a framework within which the detailed sequential aspects of behavior can be scrutinized. An experiment designed only to establish the existence of a gross relation between two variables, such as response speed and reward magnitude, ignores the many sequential properties of psychological phenomena. Examination of these properties is a significant step forward in that it provides a source of behavioral information that cannot be obtained from an analysis of average performance curves. Theories stated only in qualitative terms do not provide an adequate means for analyzing and interpreting such complex sequential phenomena.

Further, the absence of precise systematization often leads to pseudo-derivations from the theory; that is, derivations which require assumptions that are not part of the original theory. Some people claim to be unconcerned with whether the predictions tested by an experiment follow in a strictly logical sense from basic postulates. They maintain that the essential act is the making of the prediction, not its derivation from fundamental theory. The reply to this point of view seems obvious. The inability of a theory to yield significant predictions without additional ad hoc assumptions is an indication that the theory does not provide an objective analysis of behavior. An important function of the mathematical model is to clarify this aspect of a theory. Of course, many models can stem from the same fundamental theory. The important factor is whether the theory will yield at least one well-defined model in a non-arbitrary manner. The attempt to specify a model will in itself require an exact characterization of the theory and will frequently reveal unstated assumptions.

Experimental Situation

The psychophysical experiment that we shall analyze was conducted by Kinchla (1962), and he has kindly given us permission to present some of his data. He employed a forced-choice visual detection situation involving a series of over 800 discrete trials; we shall only consider data from a subset of 600 trials. Two areas were outlined on a uniformly illuminated milk glass screen. Each trial began with an auditory signal. During the auditory signal one of the following events occurred. (a) A fixed increment in radiant intensity occurred in one of the two areas of the visual display. A trial will be termed a T_1 or T_2 trial depending upon which of the two signal areas had an increment in illumination. (b) No change in the radiant character of either signal area occurred. Such blank trials will be denoted T_0 .

Subjects were instructed that a change would occur in one of the two areas on each trial. Following the auditory signal the subject was required to make either an A_1 or A_2 response to indicate which area he believed had changed in brightness; thus, the subject was forced to respond on every trial regardless of how confident he was of his choice. In this particular study by Kinchla, the subject was given no information at the end of the trial as to whether his response was correct. In summary, on a given trial one of three events occurred (T_1, T_2, T_0), the subject made either an A_1 or A_2 response, and a short time later the next trial began.

For a fixed signal intensity the experimenter has the option of specifying a schedule for presenting the T_i events. Kinchla selected a simple probabilistic procedure where the likelihood of presenting T_i on trial n was constant over all trials and independent of preceding responses and events; i.e., $\Pr(T_{i,n}) = \pi_i$ where $\pi_1 + \pi_2 + \pi_0 = 1$. Two groups of subjects were run. For Group I, $\pi_1 = \pi_2 = .4$ and $\pi_0 = .2$. For Group II, $\pi_1 = \pi_0 = .2$ and $\pi_2 = .6$.

Model

The model that will be used to describe Kinchla's experiment is a generalization of stimulus sampling concepts as originally formulated by Estes (1950). Only those axioms that are relevant to the experiment will be presented. The reader interested in a more comprehensive formulation of stimulus sampling theory is referred to Estes (1959), Estes and Suppes (1959), or Atkinson and Estes (1962); for a discussion of signal detection within the framework of stimulus sampling theory see Atkinson (1961).

In this paper, the stimulus situation is represented in terms of two <u>sensory elements</u> s_1 and s_2 and a set S^* of stimulus elements associated with background stimulation. These stimulus elements are theoretical constructs to which we assign certain properties. Although it is sometimes convenient and suggestive to speak in such terms, one should not assume that the stimulus elements are to be identified with any simple neurophysiological unit, as, for example, receptor cells. At the present stage of theory construction, we mean to assume only that certain properties of the set-theoretical model represent certain properties of the process of stimulation. If these assumptions prove to be adequately substantiated when the model is tested against a wide range of behavioral data, then it will be in order to look for neurophysiological variables that might underlie the correspondence.

On every trial the subject samples a single element from the background set S^* and may or may not sample one of the sensory elements. If the s_1 sensory element is sampled an A_1 occurs; if s_2 is sampled, A_2 occurs. If neither sensory element is sampled the subject makes the response to which the background element is conditioned. Conditioning of elements in S^* may change from trial to trial via a simple learning process. As will become evident in the statement of the axioms, we have a Fechner-type threshold model that interacts with a learning process to generate the subjects' protocol of responses.

The axioms will be formulated verbally. It is not difficult to state them in mathematical form, but for our purposes this will not be necessary. The first group of axioms deals with the sampling of stimulus elements, the second group with the conditioning process, and the third group with responses.

Stimulus Axioms

S1. If $T_i (i = 1,2)$ occurs, then sensory element s_i will be sampled with probability h.

S2. If T_0 occurs, then neither s_1 nor s_2 will be sampled.

S3. Exactly one element is sampled from set S^* on every trial. Given the set S^* of N elements, the probability of sampling a particular element is $\frac{1}{N}$, independently of the trial number and preceding events.

Conditioning Axioms

C1. On every trial each element in S^* is conditioned to either A_1 or A_2.

C2. If $s_i (i = 1,2)$ is sampled on trial n, then with probability c' the element sampled from S^* on trial n becomes conditioned to A_i at the start of trial $n + 1$.

C3. If neither s_1 nor s_2 are sampled, then with probability c the element sampled from S^* on trial n becomes conditioned with equal likelihood to either A_1 or A_2 at the start of trial $n + 1$.

Response Axioms

R1. If sensory element s_i is sampled, then the A_i response will occur.

R2. If neither sensory element is sampled, then the response to which the sampled element from set S^* is conditioned will occur.

Predicted and Observed Quantities

We begin our analysis of the model by deriving an expression for the proportion of elements in set S^* conditioned to A_1 at the start of trial n; this quantity will be denoted as p_n. Once an expression for p_n has been obtained, we immediately can write an equation for the probability of response A_i given event T_i on trial n. The expressions are obtained directly by applying axioms R1 and R2 and are as follows:

$$\Pr(A_{1,n}|T_{1,n}) = h + (1 - h)p_n \tag{1a}$$

$$\Pr(A_{2,n}|T_{2,n}) = h + (1 - h)(1 - p_n) \tag{1b}$$

$$\Pr(A_{1,n}|T_{0,n}) = p_n \tag{1c}$$

To obtain an expression for p_n, it is helpful to proceed in a series of steps. First assume that on trial n a T_1 event occurred and an element conditioned to A_1 was sampled from set S^*; the likelihood of this event is $\pi_1 p_n$. Given these events on trial n, the possible changes that can occur in p_n are specified by the tree in Figure 1. On the upper branch an s_1 sensory element is sampled with probability h (see Axiom S1) and by Axiom C2 the element sampled from S^* remains conditioned to A_1; hence on this branch $p_{n+1} = p_n$. On the lower branch neither sensory element is sampled with probability $1 - h$. By Axiom C3 there is a probability $\frac{c}{2}$ that the element sampled from S^* will become conditioned to A_2 (and hence on this branch $p_{n+1} = p_n - \frac{1}{N}$)

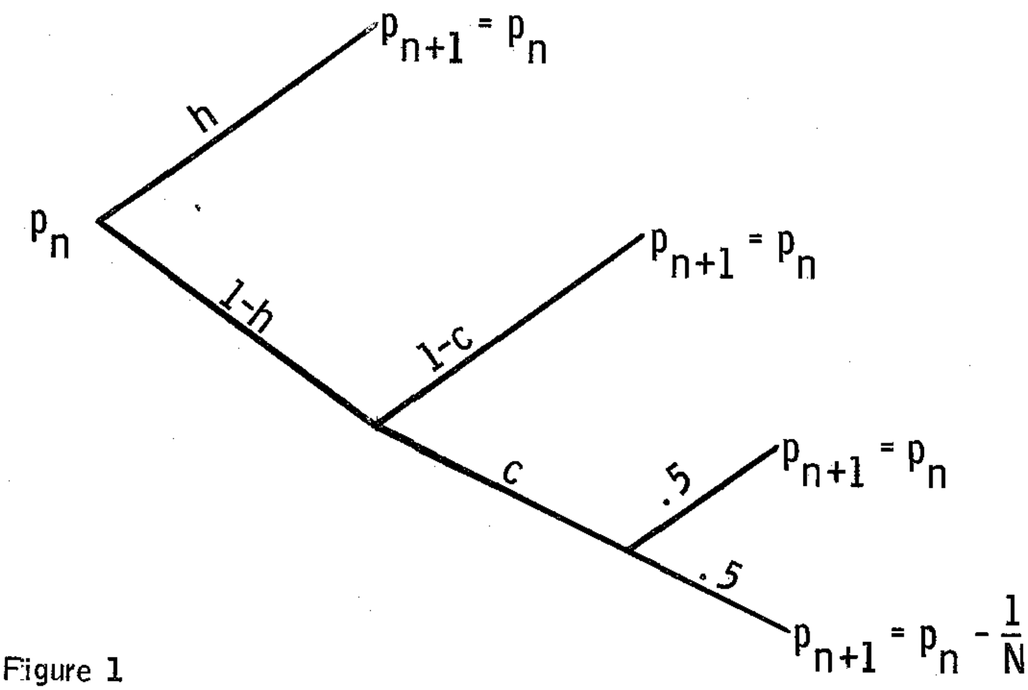

Figure 1

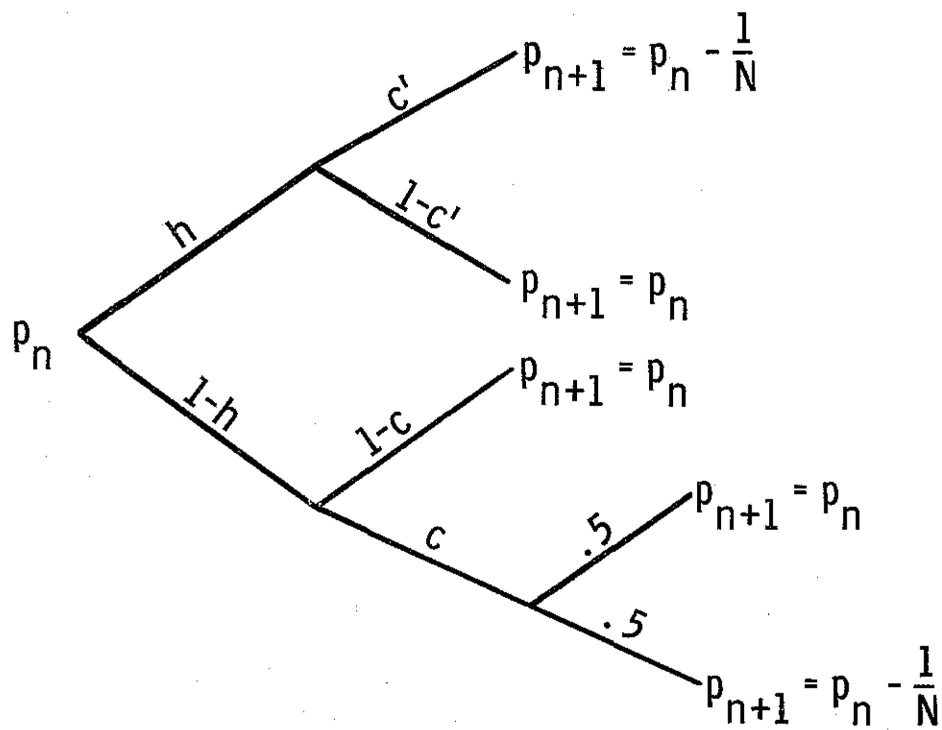

Figure 2

and a probability $1 - c + \frac{c}{2}$ that the element remains conditioned to A_1 (whence $p_{n+1} = p_n$) . Thus with probability $\pi_1 p_n$ the value of

$$p_{n+1} = h\, p_n + (1 - h)\left[(1 - \tfrac{c}{2})p_n + \tfrac{c}{2}(p_n - \tfrac{1}{N})\right]$$

or more simply $p_{n+1} = p_n - \frac{1}{N}(1 - h)\frac{c}{2}$.

Similarly, if on trial n a T_2 occurs and an element conditioned to A_1 is sampled from S^* (the probability of the joint event being $\pi_2 p_n$) then the possible changes in p_n are given by the tree in Figure 2. On the upper branch as s_2 sensory element is sampled and by Axiom C2 the element sampled from set S^* will become conditioned with probability c' to A_2 , whence $p_{n+1} = p_n - \frac{1}{N}$. The lower branches of the tree are derived by application of Axiom C3. Hence with probability $\pi_2 p_n$ the value of p_{n+1} is

$$\left[p_n - \tfrac{1}{N}\right]\left[hc' + (1 - h)\,\tfrac{c}{2}\right] + p_n\left[h(1 - c') + (1 - h)(1 - \tfrac{c}{2})\right]$$

or more simply $p_{n+1} = p_n - \frac{1}{N}\left[hc' + (1 - h)\,\frac{c}{2}\right]$.

Following the same procedure we may derive a comparable expression for p_{n+1} given the joint occurrence of event T_0 and an element sampled from S^* conditioned to A_1 ; similarly, expressions for p_{n+1} may be obtained given the joint occurrence of an element sampled from S^* conditioned to A_2 and event $T_i (i = 0, 1, 2)$. If these six results are combined weighting each by its likelihood of occurrence the following expression is obtained:

$$p_{n+1} = \pi_1 p_n\left[p_n - \frac{1}{N}(1-h)\frac{c}{2}\right] + \pi_2 p_n\left[p_n - \frac{1}{N}\left\{hc' + (1-h)\frac{c}{2}\right\}\right] + \pi_0 p_n\left[p_n - \frac{1}{N}\frac{c}{2}\right]$$

$$+ \pi_1(1-p_n)\left[p_n + \frac{1}{N}\left\{hc' + (1-h)\frac{c}{2}\right\}\right] + \pi_2(1-p_n)\left[p_n + \frac{1}{N}(1-h)\frac{c}{2}\right]$$

$$+ \pi_0(1-p_n)\left[p_n + \frac{1}{N}\frac{c}{2}\right] .$$

Collecting terms and simplifying yields a recursive expression in p_n :

$$p_{n+1} = p_n\left[1 - \frac{1}{N}(a+b)\right] + \frac{1}{N}a \quad , \tag{2}$$

where $a = \pi_1 hc' + (1-h)\frac{c}{2} + \pi_0 h\frac{c}{2}$ and $b = \pi_2 hc' + (1-h)\frac{c}{2} + \pi_0 h\frac{c}{2}$. This difference equation has the well-known solution (cf. Bush and Mosteller, 1955; Suppes and Atkinson, 1960)

$$p_n = p_\infty - (p_\infty - p_1)\left[1 - \frac{1}{N}(a+b)\right]^{n-1} \quad , \tag{3}$$

where $p_\infty = \frac{a}{a+b}$. Dividing the numerator and denominator of p_∞ by c yields the expression

$$p_\infty = \frac{\pi_1 h\psi + \frac{1}{2}(1-h) + \pi_0 h\frac{1}{2}}{(1-\pi_0)(1-h+h\psi) + \pi_0} \tag{4}$$

where $\psi = \frac{c'}{c}$. Thus, the asymptotic expression for p_n does not depend on the actual values of c' and c but only on their ratio.

With this expression at hand, we can now look at part of Kinchla's data. Figures 3 and 4 present the observed mean proportions for an A_i

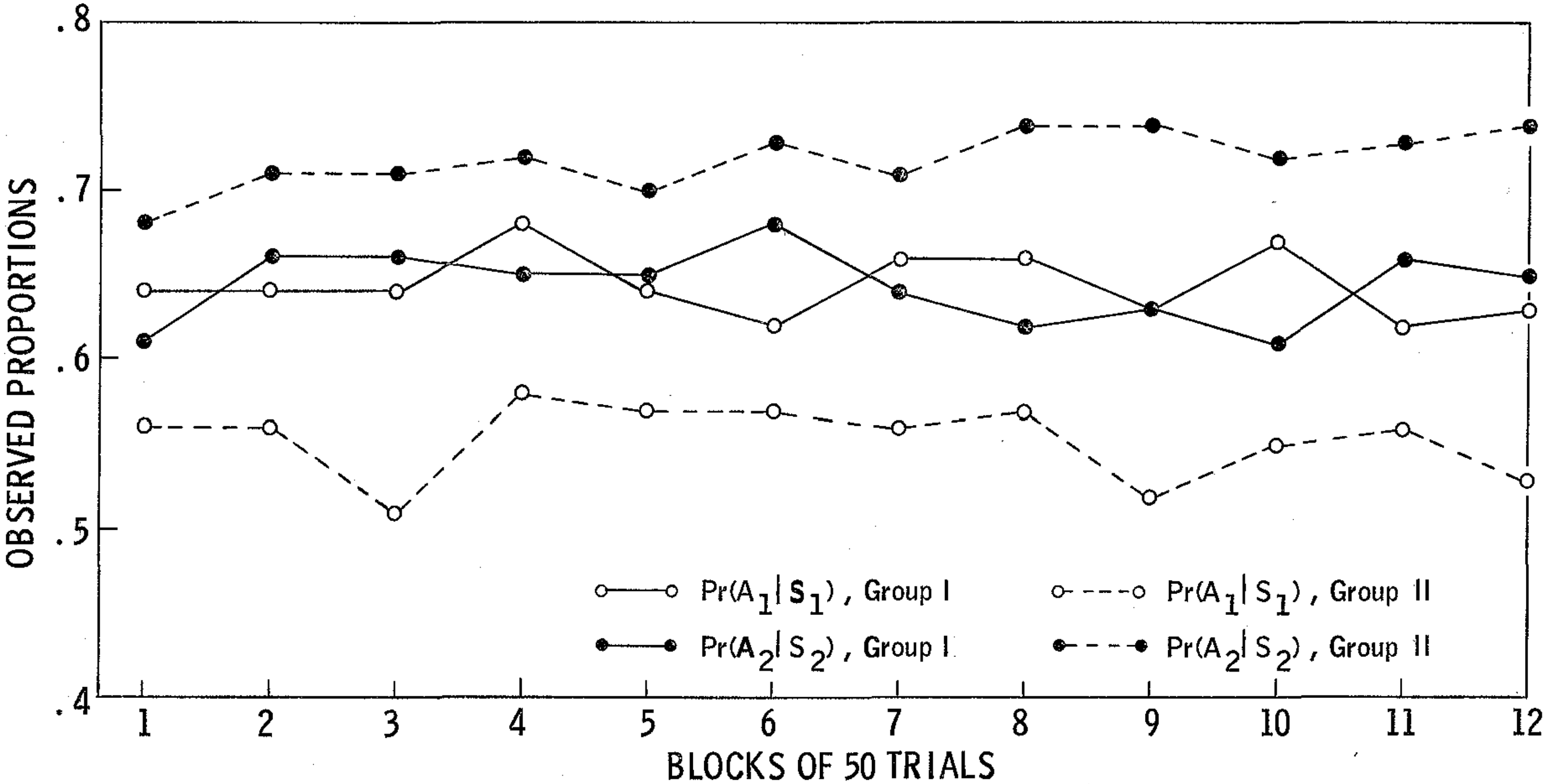

Figure 3. Observed estimates of $Pr(A_1|T_1)$ and $Pr(A_2|T_2)$ in successive 50-trial blocks.

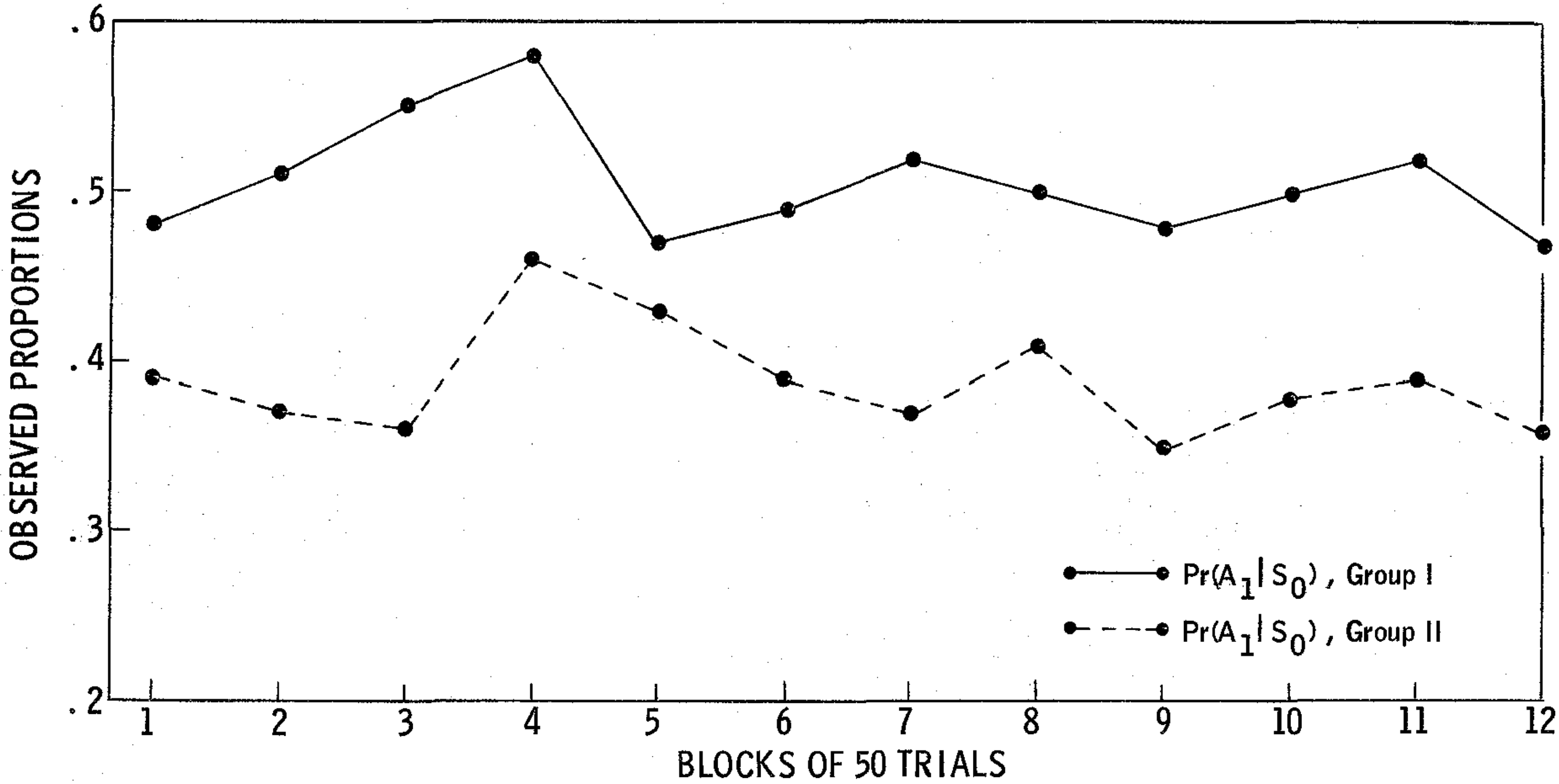

Figure 4. Observed estimates for $Pr(A_1|T_0)$ in successive 50-trial blocks.

response given a T_j event in successive blocks of 50 trials. An inspection of these curves indicates that all of the functions are quite stable over the last 400 trials. In view of this observation we have estimated asymptotic response probabilities by averaging over the last 400 trials. Table 1 presents these observed values for the two groups. The corresponding asymptotic proportions are specified in terms of Eq. 1 and Eq. 4 and are simply

$$\lim_{n \to \infty} \Pr(A_{1,n} | T_{1,n}) = h + (1 - h)p_{\infty} \tag{5a}$$

$$\lim_{n \to \infty} \Pr(A_{2,n} | T_{2,n}) = h + (1 - h)(1 - p_{\infty}) \tag{5b}$$

$$\lim_{n \to \infty} \Pr(A_{1,n} | T_{0,n}) = p_{\infty} \tag{5c}$$

In order to generate asymptotic predictions we need values for h and ψ. We first note by inspection of Eq. 4 that $p_{\infty} = \frac{1}{2}$ for Group I; in fact whenever $\pi_1 = \pi_2$ we have $p_{\infty} = \frac{1}{2}$, independent of the value of ψ. Hence, taking the observed asymptotic value for $\Pr(A_1 | T_1)$ in Group I (i.e., .645) and setting it equal to $h + (1 - h)\frac{1}{2}$ yields an estimate of $h = .289$. The physical stimuli and the increments in radiant intensity are the same for both experimental groups and therefore we would require an estimate of h obtained from Group I to be applicable to Group II. In order to estimate ψ, we take the observed asymptotic value of $\Pr(A_1 | T_0)$ in Group II and set it equal to Eq. 4 with $h = .289$, $\pi_1 = \pi_0 = .2$ and $\pi_2 = .6$; solving for ψ we obtain $\hat{\psi} = 2.8$. Using these estimates of h and ψ and Eqs. 4 and 5 yields the asymptotic predictions given in Table 1.

Over all the equations give an excellent account of these particular response measures. However, the model provides a much richer analysis of the experiment than the above results indicate. For as we have said before, the model predicts not only average performance but also detailed sequential phenomena. In terms of our axioms, sequential effects are produced by the trial-to-trial fluctuations that occur in the conditioning of elements in set S^*; such fluctuations, of course, can take place on any trial and are not restricted to pre-asymptotic data. For example, even at asymptote the likelihood of making a correct response to a T_1 event depends in a very definite way on whether an A_1 or A_2 response occurred on the preceding trial.

It should be emphasized that one of the important contributions of mathematics to behavior theory has been to provide a framework within which sequential phenomena can be analyzed. Prior to the development of mathematical models relatively little attention was given to trial-to-trial events; at the present time, for many experimental problems (especially in learning) such phenomena are viewed as the most basic feature of the data. To indicate the type of sequential predictions that can be obtained, consider the probability of an A_1 response on a T_1 trial given the various trial types and responses that can occur on the preceding trial; i.e.,

$$Pr(A_{1,n+1} | T_{1,n+1} A_{i,n} T_{j,n})$$

where $i = 1, 2$ and $j = 0, 1, 2$. Explicit expressions for these quantitites can be derived from the axioms. The actual derivations are quite lengthy and will not be presented here; the reader interested in the mathematical techniques involved should consult Atkinson and Estes (1962). Also, for purposes of this paper, the analysis of sequential statistics will be confined to asymptotic data. Therefore, only theoretical expressions for $\lim_{n \to \infty} \Pr(A_{1,n+1} | T_{1,n+1} A_{i,n} T_{j,n})$ will be given, and to simplify notation they will be written as $\Pr(A_1 | T_1 A_i T_j)$. The expressions for these quantities are as follows:

$$\Pr(A_1 | T_1 A_1 T_1) = \frac{[h + (1 - h)\delta]p_\infty + (1 - p_\infty)h\gamma'}{NX} + \frac{(N - 1)X}{N} \tag{6a}$$

$$\Pr(A_1 | T_1 A_2 T_1) = \frac{(1 - h)\delta'(1 - p_\infty)}{N(1 - X)} + \frac{(N - 1)X}{N} \tag{6b}$$

$$\Pr(A_1 | T_1 A_2 T_2) = \frac{h\gamma p_\infty + [h^2 + (1 - h)\delta'](1 - p_\infty)}{NY} + \frac{(N - 1)X}{N} \tag{6c}$$

$$\Pr(A_1 | T_1 A_1 T_2) = \frac{(1 - h)\delta p_\infty}{N(1 - Y)} + \frac{(N - 1)X}{N} \tag{6d}$$

$$\Pr(A_1 | T_1 A_1 T_0) = \frac{\delta}{N} + \frac{(N - 1)X}{N} \tag{6e}$$

$$\Pr(A_1 | T_1 A_2 T_0) = \frac{\delta'}{N} + \frac{(N - 1)X}{N} \tag{6f}$$

where $\gamma = c'h + (1 - c')$, $\gamma' = c' + (1 - c')h$, $\delta = \frac{c}{2} h + (1 - \frac{c}{2})$, $\delta' = \frac{c}{2} + (1 - \frac{c}{2})h$, $X = h + (1 - h)p_\infty$, and $Y = h + (1 - h)(1 - p_\infty)$. It is interesting to note that the asymptotic expressions for

$\lim \Pr(A_{i,n}|T_{j,n})$ depend only on h and ψ, whereas the quantities in Eq. 6 are functions of all four parameters N, c, c' and h. Comparable sets of equations can be written for $\Pr(A_2|T_2A_iT_j)$ and $\Pr(A_1|T_0A_iT_j)$. We return to this point later.

The expressions in Eq. 6 are rather formidable, but numerical predictions can be easily calculated once values for the parameters have been obtained. Further, independent of the parameter values, certain relations among the sequential probabilities can be specified. As a simple example of such a relation, it can be shown that $\Pr(A_1|T_1A_1T_0) \geq \Pr(A_1|T_1A_2T_0)$ for any stimulus schedule and any set of parameter values. To see this, simply substract Eq. 6f from Eq. 6c and note that $\delta \geq \delta'$.

In Table 2 the observed values for $\Pr(A_i|T_jA_kT_\ell)$ are presented as reported by Kinchla. Estimates of these conditional probabilities were computed for individual subjects using the data over the last 400 trials; the average of these individual estimates are the quantities given in the table. Each entry is based on 24 subjects.

In order to generate theoretical predictions for the observed entries in Table 2, values for N, c, c' and h are needed. Of course, estimates of h and $\psi = \frac{c'}{c}$ already have been made for this set of data, and therefore it is only necessary to estimate N and either c or c'. We obtain our estimates of N and c by a least squares method; i.e., we select a value of N and c (where $c' = c\hat{\psi}$) so that the sum of squared deviations between the 36 observed values in Table 2 and the corresponding theoretical quantities is minimized. The theoretical quantities for $\Pr(A_1|T_1A_iT_j)$ are computed from Eq. 6; theoretical

Table 2

Predicted and Observed Sequential Response Probabilities at Asymptote

	Group I		Group II	
	Observed	Predicted	Observed	Predicted
$Pr(A_2 \mid T_2A_1T_1)$	.57	.58	.59	.64
$Pr(A_2 \mid T_2A_2T_1)$	.65	.69	.70	.76
$Pr(A_2 \mid T_2A_2T_2)$	.71	.71	.79	.77
$Pr(A_2 \mid T_2A_1T_2)$	.61	.59	.69	.66
$Pr(A_2 \mid T_2A_1T_0)$	.54	.59	.68	.66
$Pr(A_2 \mid T_2A_2T_0)$	.66	.70	.71	.76
$Pr(A_1 \mid T_1A_1T_1)$	.73	.71	.70	.65
$Pr(A_1 \mid T_1A_2T_1)$	.62	.59	.59	.52
$Pr(A_1 \mid T_1A_2T_2)$	.53	.58	.53	.51
$Pr(A_1 \mid T_1A_1T_2)$	.66	.70	.64	.64
$Pr(A_1 \mid T_1A_1T_0)$	.72	.70	.61	.63
$Pr(A_1 \mid T_1A_2T_0)$	.61	.59	.48	.52
$Pr(A_2 \mid T_0A_1T_1)$	.38	.40	.47	.49
$Pr(A_2 \mid T_0A_2T_1)$	.56	.58	.59	.66
$Pr(A_2 \mid T_0A_2T_2)$	.64	.60	.67	.68
$Pr(A_2 \mid T_0A_1T_2)$	.47	.42	.51	.51
$Pr(A_2 \mid T_0A_1T_0)$	.47	.42	.50	.51
$Pr(A_2 \mid T_0A_2T_0)$	.60	.58	.65	.66

expressions for $Pr(A_2|T_2A_iT_j)$ and $Pr(A_2|T_0A_iT_j)$ have not been presented in this paper but are of the same general form as those given in Eq. 6.

Using this technique, estimates of the parameters are as follows:

$$\begin{aligned} N &= 4.23 \\ c' &= 1.00 \\ c &= .357 \\ h &= .289 \end{aligned} \tag{7}$$

The predictions corresponding to these parameter values are presented in Table 2. When one considers that only four of the possible 36 degrees of freedom represented in Table 2 have been utilized in estimating parameters, the close correspondence between theoretical and observed quantities in Table 2 may be interpreted as giving considerable support to the assumptions of the model. Of course, for any given subject, N must be an integer. The fact that our estimation procedure yielded a non-integral value may signify that N varies somewhat between subjects, or it may reflect some contamination of the data by sources of experimental error not represented in the model. To answer these questions a more detailed analysis of the data would be necessary in which estimates of the parameter values are made for individual subjects. Such analyses are too lengthy to discuss here.

Discussion

No model can be expected to give a perfect account of fallible data arising from real experiments as distinguished from the idealized experiment to which the model should apply strictly. Consequently, it is difficult to know how to evaluate the goodness-of-fit of theoretical to observed values. In practice investigators usually proceed on a largely intuitive basis, evaluating the fit in a given instance against that which it appears reasonable to expect in view of what is known about the precision of experimental control and measurement. Statistical tests of goodness-of-fit are sometimes possible (discussions of some tests that may be used in conjunction with stimulus sampling models are given by Suppes and Atkinson, 1960); however, statistical tests are not entirely satisfactory taken by themselves, for a sufficiently precise test will often indicate significant differences between theoretical and observed values even in cases where the agreement is as close as could reasonably be expected. Generally, once a degree of descriptive accuracy has been attained that appears satisfactory to investigators familiar with the given research area, further progress must come largely via differential tests of alternative models.

The problems involved in making differential tests among various models for signal detection are beyond the scope of this paper. However, it is clear that alternative models can be formulated that deserve careful analysis. For example, in this paper we have examined a very special Markovian conditioning process defined on the background stimuli; it would be important to determine whether other formulations of the learning process, such as those developed by Bush and Mosteller (1955), would provide as good or even better fits. Also, it would be valuable to consider variations in the scheme for sampling sensory elements along lines developed by Luce (1959) and Restle (1961).

Independent of further analysis, it appears that this particular model for signal detection provides an impressive account of average response proportions and of various complex sequential events. Some readers may object to the model and argue that the axioms are unrealistic from a physiological viewpoint or in terms of a cognitive analysis. Such objections are important if they generate new ideas. However, in the absence of concrete suggestions that lead to testable models it is doubtful whether this type of criticism is meaningful. For on that basis one might just as well object to the kinetic theory of gases because it assumes that gas molecules behave as perfectly rigid spheres or to classical hydrodynamics because it postulates complete continuity of fluids. Certainly there is no single true theory. Rather there is an infinite number of hypotheses and theories that can explain an array of phenomena, and we judge a particular theory not by some intuitive notion of reality but in terms of the theory's ability to account for the facts at hand and to generate new predictions.

In our concluding remarks, it would be nice if we could refer to a list of criteria or a decision rule that would evaluate the approach taken in this paper and tell us whether this specific development or related mathematical models are of genuine value in analyzing psychological phenomena. Of course, such decision procedures do not exist. Only the perspective gained by refinement and extension of these models with empirical verification at critical stages will permit us to make such an evaluation. Certainly within the last decade many behavioral phenomena have been examined with reference to one or more mathematical models, and there is no doubt that these analyses have led to a deeper understanding of the empirical findings. In addition, many new lines of experimental research have been initiated by work on mathematical models. Despite these developments some behavioral scientists maintain that psychology has not yet reached a stage where mathematical analysis is appropriate; still others argue that the data of psychology are inherently different from those of the natural sciences and defy any type of rigorous systematization. There is no conclusive answer to these criticisms. Similar objections were raised against mathematical physics as recently as the late 19th century, and only the brilliant success of the approach silenced opposition. A convincing argument is yet to be made for the possibility that mathematical psychology will not enjoy similar success.

References

Atkinson, R.C. A variable threshold model for signal detection. Technical Report No. 42, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford Univ., 1961.

Atkinson, R.C., and Estes, W.K. Stimulus sampling theory. In Bush, R.R., Galanter, E. and Luce, R.D. (Eds.), _Handbook of mathematical psychology_. New York: Wiley (in preparation).

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

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

Estes, W.K. The statistical approach to learning theory. In Koch, S. (Ed.), _Psychology: A study of a science_ (New York: McGraw-Hill, 1959), II, 383-491.

Estes, W.K. and Suppes, P. Foundations of statistical learning theory, II: The stimulus sampling model for simple learning. Technical Report No. 26, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford Univ., 1959.

Kinchla, R.A. Learned factors in visual discrimination. Unpublished doctor's dissertation, University of California, Los Angeles, 1962.

Luce, R.D. _Individual choice behavior: A theoretical analysis_. New York: Wiley, 1959.

Restle, F. _Psychology of judgment and choice: A theoretical essay_. New York: Wiley, 1961.

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