SEQUENTIAL PHENOMENA IN PSYCHOPHYSICAL JUDGMENTS:

A THEORETICAL ANALYSIS*/

by

R. C. Atkinson

E. C. Carterette

R. A. Kinchla

TECHNICAL REPORT NO. 46

April 20, 1962

PSYCHOLOGY SERIES

---

*/ This paper is to be presented at the International Symposium on Information Theory, Brussels, September 3 to 7, 1962 and will be printed beforehand in a special issue of The Transactions of the Institute of Radio Engineers. The research was supported in part by the National Institute of Mental Health under Contract M-5184.

SEQUENTIAL PHENOMENA IN PSYCHOPHYSICAL JUDGMENTS: A THEORETICAL ANALYSIS

R. C. Atkinson
Stanford University
Stanford, California

E. C. Carterette
University of California
Los Angeles, California

R. A. Kinchla
Ames Research Center
Moffett Field, California

Summary

This paper deals with an analysis of psychophysical detection experiments designed to assess the limit of a human observer's level of sensitivity. A mathematical theory of the detection process is introduced that, in contrast to previous theories, provides an analysis of the sequential effects observed in psychophysical data. Two variations of the detection task are considered: information feedback and no-information feedback. In the feedback situation the subject is given information concerning the correctness of his responses, whereas in the no-feedback situation he is not. Data from a visual detection experiment with no-information feedback, and from an auditory detection experiment with information feedback are analyzed in terms of the theory. Finally, some general results are derived concerning the relationship between performance in the feedback situation and the no-feedback situation.

Introduction

This paper presents an analysis of the process by which a human observer detects the occurrence of very weak signals. The theoretical formulation that we offer should apply to signals received via any sensory mode, but our discussion will be restricted to visual and auditory stimuli. Furthermore, the analysis is behavioral rather than physiological since it deals with the subjects' overt responses rather than with biochemical or neurophysiological activity.

A methodology for assessing the limits of a subject's sensitivity to external stimuli based on phenomenal reports was developed quite early (Fechner, 1860) and has remained relatively unchanged since that time. Most simply, these methods offered a means for determining the probability of a "detection" for various signal intensities. Early investigators often interpreted the subject's phenomenal report quite directly; i.e., a reported detection implied that the signal was above the subject's limit of sensitivity and a report of no detection implied that it was below this limit. The limit, or threshold as it has often been called, was viewed as varying randomly in time about a fixed mean value. Therefore, the threshold was defined statistically as that signal intensity reported by the subject on half of the occasions on which it was presented. More recently, alternative interpretations of the subject's performance have been proposed (e.g., see Blackwell[1] and Swets, Tanner, and Birdsall[2]). These proposals all view the subject as utilizing more than the immediate sensory information to determine his response on each trial. However, these newer approaches are still traditional in at least one major respect: they represent the detection process as fixed over long series of trials. This static conception of psychophysical phenomena is surprising in view of the sequential effects that are apparent in the trial-to-trial data. Investigators as far back as Fechner[3] have noted that the subject's response tendency on one trial is markedly influenced by the stimuli and responses that occur on preceding trials. Most investigators either have ignored these sequential effects or treated them as experimental artifacts, to be minimized by randomization, counterbalancing, trial spacing, or by use of trained subjects. In this paper sequential effects will be considered as an important aspect of the subject's performance; furthermore it is our contention that consideration of these effects provides a valuable insight into the character of the detection process. Specifically, we deal with an analysis of sequential statistics in two types of detection situations; one situation involves information feedback on each trial, the other does not.

The type of psychophysical situation that we analyze is a two-response, forced-choice detection task. On each trial two temporal, or spatial, intervals are defined and the subject is instructed to report which of these two intervals contained a signal. It is a forced-choice task in that on each trial the subject must identify one of the two intervals as containing a signal even if he is uncertain as to what occurred. The following notation will be used to identify each trial:

$T_{i,n}$ = the presentation of a signal in interval i on trial n $(i=1, 2)$; or the presentation of a signal in neither interval $(i=0)$.

$A_{j,n}$ = the subject's selection of interval j $(j=1, 2)$ as the interval containing the signal on trial n .

$E_{k,n}$ = the occurrence of an information event at the end of trial n which informs the subject that the signal has occurred in interval k $(k=1, 2)$; or no information at the conclusion of trial n $(k=0)$.

Using this notation, each trial may be described by an ordered triple (T_i,A_j,E_k) .

As indicated above, the two variations of the detection task that we analyze in this paper are information feedback and no-information feedback. The information condition requires that the experimenter present E_1 on a T_1 trial and E_2 on a T_2 trial; the no-information case requires that

E_0 occurs on all trials. In addition to these two cases one can also study the effects of presenting incorrect information on some trials. Carterette and Wyman[4] have investigated the influence of misinformation, and the theory we present here is applicable to their experiment. However, to simplify our discussion we shall not examine the misinformation condition.

When no information is given to the subject it seems natural on occasion to introduce a "blank" trial and note its effect on choice behavior. Hence for the no-information condition we permit T_0 trials. However, the introduction of T_0 trials in the information condition raises problems regarding the type of feedback that should be given on these trials; to avoid these special issues we restrict our analysis of the information case to situations involving only T_1 and T_2 trials. Thus for the no-information case the experimenter has the option of presenting $T_1 - E_0$, $T_2 - E_0$, or $T_0 - E_0$ on each trial. For the information case he may present either $T_1 - E_1$ or $T_2 - E_2$. In this paper we consider only simple probabilistic schedules for presenting events. For the information case we denote the probability of the two events as follows:

$$\gamma = \Pr(T_1 \ \& \ E_1)$$

$$1 - \gamma = \Pr(T_2 \ \& \ E_2) \ .$$

For the no-information case:

$$\pi_1 = \Pr(T_1 \ \& \ E_0)$$

$$\pi_2 = \Pr(T_2 \ \& \ E_0)$$

$$\pi_0 = \Pr(T_0 \ \& \ E_0) \ ,$$

where $\pi_1 + \pi_2 + \pi_0 = 1$.

Theory

Before we turn to a discussion of the theory on which our analysis is based, a few general remarks will be useful. All psychological theories of signal detection incorporate two distinct processes: an activation process and a decision process. The activation process specifies the relation between external stimulus events and hypothesized sensory states of the subject. The decision process specifies the relation between the sensory states and the observable response of the subject. For example, the model proposed by Blackwell[1] may be interpreted in terms of these two processes. Two sensory states, "true detection" and "no detection", are defined and the activation process is characterized by specifying the probability that one of these two sensory states occurs for a given signal intensity. The decision process is characterized by specifying the probability of the subject's response for each of these two sensory states. In Blackwell's model the subject always makes the correct response given a "true detection", but guesses one response or the other with some fixed probability when the "no detection" state occurs. Other models of the detection process (e.g., Swets, Tanner, and Birdsall[2]) have more complicated views of the activation and decision processes. However, all of these models are similar in one respect: the character of the activation and decision processes is viewed as fixed over long series of trials. It is this common feature that was referred to earlier as a static view of the detection process. The general theory used in our analysis was developed by Atkinson[5,6] and considers both the activation and decision processes as varying from trial to trial. However, a satisfactory treatment of the problems that we consider in this paper can be obtained by using a special case of the general theory; for this case only the decision process is viewed as dynamic.

The theoretical representation that will be used here is a generalization of stimulus sampling concepts as originally formulated by Estes[7]; a comprehensive survey of stimulus sampling theory may be found in Atkinson and Estes[8]. For purposes of this paper the stimulus situation will be represented in terms of two sensory patterns, s_1 and s_2, and a set S of stimulus patterns associated with background stimulation. These patterns 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 these patterns are to be identified with any simple neurophysiological unit such as a receptor cell. 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 a single pattern is activated from the background set S, and simultaneously one of the sensory patterns may or may not be activated. If the s_1 sensory pattern is activated A_1 occurs; if s_2 is activated A_2 occurs. If neither sensory pattern is activated the subject makes the response to which the background pattern is conditioned. Conditioning of patterns in S may change from trial to trial via a simple learning process. It is the manner in which this conditioning process is conceptualized that distinguishes the information situation from the no-information situation. In the feedback situation the information event itself controls the conditioning process; without feedback the conditioning process is controlled by the sensory pattern activated on each trial. This distinction will become clear after consideration of the axioms. The axioms will be formulated verbally; it is not difficult to state them in mathematically exact form, but for present purposes this is not necessary. The axioms fall into three groups: the first group defines the activation process, the second group defines the decision process, and the third group defines the manner in which the conditioning of background elements occurs. Two sets of conditioning axioms will be stated: one set is applicable to the information case, and the other to the no-information case.

Activation Axioms

A1. If $T_i (i=1, 2)$ occurs, then sensory pattern s_i will be activated with probability h (with probability $1-h$ neither s_1 nor s_2 will be activated).

A2. If T_0 occurs, then neither s_1 nor s_2 will be activated.

A3. Exactly one pattern is activated from set S on every trial. Given the set S of N patterns, the probability of activating a particular pattern is $1/N$.

Response Axioms

R1. If sensory pattern s_i is activated, then the A_i response will occur.

R2. If neither sensory pattern is activated, then the response to which the pattern activated from set S is conditioned will occur.

Conditioning Axioms: No Information Feedback

C1. On every trial each pattern in S is conditioned to either A_1 or A_2.

C2. If s_i $(i=1, 2)$ is activated on trial n, then with probability c' the pattern activated from S on the trial becomes conditioned to A_i at the end of trial n.

C3. If neither s_1 nor s_2 are activated on trial n, then with probability c the pattern activated from S on the trial becomes conditioned with equal likelihood to either A_1 or A_2 at the end of trial n.

Conditioning Axioms: Information Feedback

C1. On every trial each pattern in S is conditioned to either A_1 or A_2.

C2. The pattern activated from S on each trial becomes conditioned with probability θ to the A_i response if E_i occurs on that trial; if it is already conditioned to that response, it remains so.

Thus the information case differs from the no-information case in that in the former the feedback, $E_{i,n}$, is the reinforcing event on trial n, whereas in the no-feedback case the patterns activated on trial n determine the conditioning process.

The symbol p_n will be used to denote the proportion of elements in set S conditioned to A_1 at the start of trial n. The expression for p_n will differ for the information and the no-information conditions. However, once the expression for p_n has been derived (for either the information or no-information case) the equations for the probability of response A_i given event T_i on trial n may be written immediately. These expressions are obtained by the application of axioms R1 and R2 and are as follows:

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

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

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

It will be recalled that our discussion is restricted to cases where T_0 trials only occur when there is no information feedback; consequently Eq. 1c will only be applicable to the no-feedback case.

Application to No-Feedback Data

In this section we shall evaluate data from a detection study by Kinchla[9] in which no-information feedback was given to the subject. A two-response, forced-choice, visual detection task was used and each subject was run for a series of over 800 trials; we shall only consider data from the last 400 trials. Two areas were outlined on a uniformly illuminated milk glass screen and the beginning of each trial was indicated by an auditory signal. During the auditory signal one of three possible events occurred: a fixed increment in radiant intensity occurred on one of the two areas of the visual display, or no change occurred in either area. 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; trials on which no change occurred will be termed T_0 trials. As indicated earlier, the probability of a T_i trial will be denoted π_i. 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 (press one of two keys) to indicate which area he felt had changed in brightness. No information was given him about the correctness of his response.

We shall begin our analysis of this study by considering the expression for p_n. This expression may be derived from the model by applying the conditioning axioms for the no-feedback case. Since detailed derivations of the relevant expressions for this case are available elsewhere (Atkinson[5]) these derivations will not be repeated here. However, the techniques used in the derivations are analogous to those used in the information case which is discussed later in this paper. A direct application of the conditioning axioms and subsequent simplification yields the following expression for p_n:

$$p_n = p_\infty - (p_\infty - p_1)\left[1 - \frac{1}{N}(a+b)\right]^{n-1} \tag{2}$$

where

$$a = \pi_1 hc' + (1-h)\frac{c}{2} + \pi_0 h\frac{c}{2},$$

$$b = \pi_2 hc' + (1-h)\frac{c}{2} + \pi_0 h\frac{c}{2}.$$

And

$$p_\infty = \frac{\pi_1 h\psi + \frac{1}{2}(1-h) + \pi_0 h\frac{1}{2}}{(1-\pi_0)(1-h+h\psi)+\pi_0}, \qquad (3)$$

where $\psi = c'/c$. It is interesting to note that the asymptotic expression, p_∞ , does not depend on the absolute values of c' and c but on their ratio, ψ . Throughout the remainder of this paper we shall only present mathematical results for the limiting case in which $n \to \infty$. The reason is that all the data we consider in this paper was obtained after the subject had already been run for a large number of trials. Hence, the data can best be interpreted in terms of the asymptotic form of the theory.

Using Eq. 3 we shall now consider one aspect of the data from Kinchla's study. Two groups of 24 subjects were run: Group I employed a presentation schedule where $\pi_1 = \pi_2 = .4$, and $\pi_0 = .2$; for Group II, $\pi_1 = .2$, $\pi_2 = .6$, and $\pi_0 = .2$. The average proportion of A_j responses made on T_j trials over the last 400 trials was computed for each group of subjects; the values are given in Table 1. The corresponding asymptotic proportions are specified in terms of Eqs. 1 and 3, and are simply:

$$\lim_{n\to\infty} \Pr(A_{1,n}|T_{1,n}) = h + (1-h)p_\infty \qquad (4a)$$

$$\lim_{n\to\infty} \Pr(A_{2,n}|T_{2,n}) = h + (1-h)(1-p_\infty) \qquad (4b)$$

$$\lim_{n\to\infty} \Pr(A_{1,n}|T_{0,n}) = p_\infty . \qquad (4c)$$

Consideration of Eq. 3 reveals that $p_\infty = 1/2$ if $\pi_1 = \pi_2$; thus, $p_\infty = 1/2$ for Group I. By setting the observed asymptotic value for $\Pr(A_1|T_1)$ in Group I (i.e., .645) equal to $h + (1-h)1/2$ an estimate of $h = .289$ was obtained. Since there was no relevant systematic difference in the two groups' experimental situation this estimate of h is appropriate for both groups. An estimate of ψ was obtained by setting the observed value of $\Pr(A_1|T_0)$ in Group II equal to Eq. 3 with $h = .289$, $\pi_1 = \pi_0 = .2$ and $\pi_2 = .6$; this method yielded an estimate of $\psi = 2.8$. Using these estimates of h and ψ , Eqs. 3 and 4 generate the asymptotic predictions given in the top panel of Table 1. It is apparent that the model provides a reasonably close fit to this aspect of the data.

In contrast to a static theory of a signal detection the present theory provides a much deeper analysis of the experiment than indicated by the predictions summarized in the top panel of Table 1; the dynamic character of our model allows an analysis of sequential effects as well as average performance. In the model these sequential effects are produced by the trial-to-trial fluctuations that occur in the conditioning of patterns in set S.

The notation $\Pr(A_i|T_jA_kT_m)$ will be used to represent the asymptotic probability of an A_i response on a T_j trial when the previous trial had been a T_m trial on which an A_k response was made. Eqs. 5a through 5f are expressions for these quantities derived from the axioms of the model. Since the derivations are quite lengthy they will not be given here; the reader interested in the mathematical techniques involved should consult Atkinson and Estes[8].

$$\Pr(A_1|T_1A_1T_1) = \frac{[h+(1-h)\delta]p_\infty + (1-p_\infty)h\xi'}{NX} + \frac{(N-1)X}{N} \qquad (5a)$$

$$\Pr(A_1|T_1A_2T_1) = \frac{(1-h)\delta'(1-p_\infty)}{N(1-X)} + \frac{(N-1)X}{N} \qquad (5b)$$

$$\Pr(A_1|T_1A_2T_2) = \frac{h\xi p_\infty + [h^2+(1-h)\delta'](1-p_\infty)}{NY} + \frac{(N-1)X}{N} \qquad (5c)$$

$$\Pr(A_1|T_1A_1T_2) = \frac{(1-h)\delta p_\infty}{N(1-Y)} + \frac{(N-1)X}{N} \qquad (5d)$$

$$\Pr(A_1|T_1A_1T_0) = \frac{\delta}{N} + \frac{(N-1)X}{N} \qquad (5e)$$

$$\Pr(A_1|T_1A_2T_0) = \frac{\delta'}{N} + \frac{(N-1)X}{N} \qquad (5f)$$

where $\xi = c'h + (1-c')$, $\xi' = c' + (1-c')h$, $\delta = (c/2)h + (1-c/2)$, $\delta' = c/2 + (1-c/2)h$, $X = h + (1-h)p_\infty$, and $Y = h + (1-h)(1-p_\infty)$.

Comparable sets of equations can be written for $\Pr(A_2|T_2A_kT_m)$ and $\Pr(A_1|T_0A_kT_m)$ and are of the same general form as those in Eq. 5.

Two points of interest regarding these expressions should be noted. First, the average response probabilities defined in Eq. 4 depend only on h and ψ , whereas the quantities in Eq. 5 are functions of all four parameters N , c , c' , and h. Second, independently of the parameter values certain relations among the sequential probabilities can be specified; e.g., $\Pr(A_1|T_1A_1T_0) \geq \Pr(A_1|T_1A_2T_0)$ for any stimulus schedule and any set of parameter values. To see this, simply subtract Eq. 5f from Eq. 5e and note that $\delta \geq \delta'$.

In Table 1 the observed values for $\Pr(A_i|T_jA_kT_m)$ are presented; these are based on the same data as the observed values of $\Pr(A_i|T_j)$ presented in Table 1. In order to generate theoretical predictions for the observed sequential entries in Table 1, values of N , c , c' , and h are needed. Since estimates of h and $\psi = c'/c$ already have been made for this set of data, it was only necessary to estimate N and either c or c' . The predicted values in Table 1 are based on a least squares estimate of N and c' ; i.e., N and c' were chosen so as to minimize the sum of the squared deviations between the 36 observed values in Table 2 and the corresponding theoretical values. The values of the four parameters that were used to generate the predictions are as follows: $N = 4.23$, $c' = 1.00$, $c = .357$, and

Table 1

Predicted and Observed Response Probabilities in the Visual Experiment

	Group I		Group II	
	Observed	Predicted	Observed	Predicted
$Pr(A_1\|T_1)$	.645	.645	.558	.565
$Pr(A_2\|T_2)$	.643	.645	.730	.724
$Pr(A_1\|T_0)$	.494	.500	.388	.388
$Pr(A_2\|T_2\,A_1\,T_1)$	.57	.58	.59	.64
$Pr(A_2\|T_2\,A_2\,T_1)$	.65	.69	.70	.76
$Pr(A_2\|T_2\,A_2\,T_2)$	.71	.71	.79	.77
$Pr(A_2\|T_2\,A_1\,T_2)$	.61	.59	.69	.66
$Pr(A_2\|T_2\,A_1\,T_0)$	.54	.59	.68	.66
$Pr(A_2\|T_2\,A_2\,T_0)$	.66	.70	.71	.76
$Pr(A_1\|T_1\,A_1\,T_1)$	.73	.71	.70	.65
$Pr(A_1\|T_1\,A_2\,T_1)$	.62	.59	.59	.52
$Pr(A_1\|T_1\,A_2\,T_2)$	.53	.58	.53	.51
$Pr(A_1\|T_1\,A_1\,T_2)$	.66	.70	.64	.64
$Pr(A_1\|T_1\,A_1\,T_0)$	.72	.70	.61	.63
$Pr(A_1\|T_1\,A_2\,T_0)$	.61	.59	.48	.52
$Pr(A_2\|T_0\,A_1\,T_1)$	.38	.40	.47	.49
$Pr(A_2\|T_0\,A_2\,T_1)$	.56	.58	.59	.66
$Pr(A_2\|T_0\,A_2\,T_2)$	.64	.60	.67	.68
$Pr(A_2\|T_0\,A_1\,T_2)$	.47	.42	.51	.51
$Pr(A_2\|T_0\,A_1\,T_0)$	.47	.42	.50	.51
$Pr(A_2\|T_0\,A_2\,T_0)$	.60	.58	.65	.66

$h = .289$. Since only four of the 36 possible degrees of freedom represented in Table 1 have been utilized in estimating parameters, the close fit provided by the model lends considerable support to the conception of the detection process made explicit in the axioms.

Application to Feedback Data

To indicate the nature of the predictions for an information feedback problem we shall examine some data from two subjects run in a two-response, forced-choice auditory detection task. Two temporal intervals were defined on each trial by a pair of lights. A band-limited Gaussian noise (the masking stimulus) was present continuously throughout the experimental situation and on every trial one of the two intervals contained a fixed intensity, 1000 cps tone. The subject pulled one of two levers to indicate which of the two intervals he believed contained the signal. Each trial ended with the location of the signal being indicated to the subject by another set of lights. The experimental procedure is described in detail in Atkinson and Carterette[10]; that paper deals with an analysis of forced-choice and yes-no data from six subjects, each run for 350 trials per day for 30 days. The preliminary data we present here is not to be regarded as a test of the theory, but only a means of illustrating some of the predictions. A trial will be denoted T_1 or T_2 depending upon whether the first or second interval contained the 1000 cps signal, and a correct response on a T_i trial $(i = 1, 2)$ will be termed an A_i response. Thus each T_i trial concludes with an E_i event which indicates to the subject that an A_i response was the correct response on that trial. The probability of a T_1 or T_2 (γ and $1-\gamma$, respectively) was set at $1/2$ for the data that will be considered.

The first step in our analysis of the information feedback case will be to derive an expression for p_n. To do this we first note that (whatever the value of p_n) the value of p_{n+1} will be either p_n, $p_n + 1/N$, or $p_n - 1/N$; the reason for this is that (by axiom A3) if a change in the conditioning of the background pattern occurs it will only involve one of the N patterns. Thus p_{n+1} may be written as an average of these three possible values, with each value weighted by its probability of occurrence. The probability of occurrence of each of these values may be determined directly from the axioms. For example, the probability that $p_{n+1} = p_n + 1/N$ is simply the probability that T_1 occurred on trial n (γ), times the probability that the background pattern sampled on that trial is conditioned to an A_2 response $(1-p_n)$, times the probability that conditioning is effective on that trial (θ); i.e., $Pr(p_{n+1} = p_n + 1/N) = \gamma(1-p_n)\theta$. In a similar fashion $Pr(p_{n+1} = p_n - 1/N)$ may be shown to equal $(1-\gamma)p_n\theta$. Finally, since p_{n+1} must be one of three values, $Pr(p_{n+1} = p_n) = 1 - \gamma(1-p_n)\theta - (1-\gamma)p_n\theta$. Thus the following recursive expression for p_{n+1} may be written:

$$p_{n+1} = \gamma(1-p_n)\theta\,(p_n + \tfrac{1}{N}) + (1-\gamma)p_n\theta\,(p_n - \tfrac{1}{N}) + [1 - \gamma(1-p_n)\theta - (1-\gamma)p_n\theta]p_n\,. \qquad (6)$$

This recursion can be solved by standard methods (see Atkinson and Estes[8]) to yield the explicit formula

$$p_n = p_\infty - (p_\infty - p_1)(1 - \tfrac{\theta}{N})^{n-1} \qquad (7a)$$

where

$$p_\infty = \gamma\,. \qquad (7b)$$

Table 2 presents the observed values for $Pr(A_1|T_1)$ and $Pr(A_2|T_2)$. Since $\gamma = 1/2$ we have immediately (via Eq. 7) that $p_\infty = 1/2$. Knowing p_∞ and the observed value of $Pr(A_1|T_1) = .73$ we may use Eq. 4a to obtain an estimate of $h = .46$. Using this estimate of h the model predicts that $Pr(A_1|T_2) = .27$ which is quite close to the observed value of .28.

Table 2

Predicted and Observed Response Probabilities in the Auditory Experiment

	Observed	Predicted
$Pr(A_1\|T_1)$	.73	.73
$Pr(A_1\|T_2)$	.28	.27
$Pr(A_1\|T_1A_1T_1)$	.80	.78
$Pr(A_1\|T_1A_2T_1)$	.76	.75
$Pr(A_1\|T_1A_1T_2)$	.73	.71
$Pr(A_1\|T_1A_2T_2)$	.67	.68
$Pr(A_1\|T_2A_1T_1)$	.30	.32
$Pr(A_1\|T_2A_2T_1)$	.32	.29
$Pr(A_1\|T_2A_1T_2)$	.26	.25
$Pr(A_1\|T_2A_2T_2)$	.22	.22

Expressions for the sequential probabilities, $Pr(A_1|T_iA_kT_m)$, may be derived for the feedback case just as they were for the no-feedback case. Once again, since the derivations are rather lengthy, only the expressions themselves will be presented here.

$$Pr(A_1|T_1A_1T_1) = \frac{p_\infty + (1-p_\infty)h\,\{\theta + (1-\theta)h\}}{NX} + \frac{(N-1)X}{N} \qquad (8a)$$

$$Pr(A_1|T_1A_2T_1) = \frac{(1-p_\infty)(1-h)\,\{\theta + (1-\theta)h\}}{N(1-X)} + \frac{(N-1)X}{N} \qquad (8b)$$

$$\Pr(A_1|T_1A_1T_2) = \frac{p_\infty(1-h)\ (1-\theta+\theta h)}{N(1-Y)} + \frac{(N-1)X}{N} \quad (8c)$$

$$\Pr(A_1|T_1A_2T_2) = \frac{p_\infty\ h(1-\theta+\theta h)}{NY} + \frac{(N-1)X}{N} \quad (8d)$$

where, as in Eq. 5, $X = h + (1-h)p_\infty$ and $Y = h + (1-h)(1-p_\infty)$. Comparable equations can be written for $\Pr(A_1|T_2A_iT_j)$.

To generate theoretical predictions for $\Pr(A_1|T_iA_jT_k)$ estimates of θ and N are needed, in addition to our estimate of h. Once again, we obtain our estimate of N and θ by a least squares method; i.e., values of N and θ were selected that minimized the sum of the squared deviations between the observed values for $\Pr(A_1|T_iA_jT_k)$ in Table 2 and the corresponding predictions generated by Eq. 8. The values of the three parameters that were used to generate the predictions for Table 2 are $h = .46$, $\theta = .62$, and $N = 3.83$. Since only three of the possible eight degrees of freedom represented in Table 2 have been utilized the fit is reasonably good.

Comparison of the Information and No-Information Cases

So far we have examined the types of analyses that are possible in both a feedback situation and a no-feedback situation. We now turn to a comparison of these two situations in terms of our model. Experimentally an obvious way to explore the differences between these two situations would be to conduct a study in which the same subjects were run on two identical forced-choice detection tasks that differed only with respect to the presence or absence of information feedback. Either the visual or the auditory detection task discussed previously would be appropriate for such a study. To simplify the comparison we shall permit only T_1 or T_2 type trials to occur in both the information and no-information conditions and will use the notation γ and $1-\gamma$ to denote the probabilities of a T_1 and T_2 event, respectively. Thus when interpreting the equations for the no-information case we need to set $\pi_1 = \gamma$, $\pi_2 = 1-\gamma$, and $\pi_0 = 0$.

In our hypothetical study, the same subject would be used in both feedback conditions, and also the same physical signal parameters would be used throughout the experiment. Hence, it would be reasonable to assume that h and N are the same under both the information and no-information conditions. Given this assumption we note, by comparing Eqs. 3 and 7, that

$$p_n^{I} \geq p_n^{\bar{I}} \quad \text{for} \quad \gamma > \frac{1}{2}$$

$$p_n^{I} = p_n^{\bar{I}} \quad \text{for} \quad \gamma = \frac{1}{2} ,$$

where I and $\bar{I}$ refer to the information and no-information situations, respectively. Thus using Eq. 1 (which is applicable to both situations) we have for $\gamma \geq 1/2$,

$$\overset{I}{\Pr}(A_{1,n}|T_{1,n}) \geq \overset{\bar{I}}{\Pr}(A_{1,n}|T_{1,n})$$
$$\overset{I}{\Pr}(A_{2,n}|T_{2,n}) \leq \overset{\bar{I}}{\Pr}(A_{2,n}|T_{2,n}) . \quad (10)$$

The equality holds only when $\gamma = 1/2$ or when $\psi \to \infty$. With these results in mind consider the overall probability of a correct response on trial n ; namely

$$\Pr(C_n) = \gamma \Pr(A_{1,n}|T_{1,n}) + (1-\gamma) \Pr(A_{2,n}|T_{2,n}) .$$

In view of the inequalities in Eq. 10 we immediately have

$$\overset{I}{\Pr}(C_n) \geq \overset{\bar{I}}{\Pr}(C_n) , \quad (11)$$

where, again, equality holds when $\gamma = 1/2$ or when $\psi \to \infty$. Thus in terms of overall performance, the theory predicts that the subjects will tend to be correct more often in an information feedback situation than in a no-information feedback situation.

A comparison of sequential statistics in the information and no-information cases is more tedious and for purposes of this paper we examine only one prediction; namely

$$\Delta = \Pr(A_1|T_1A_1T_1) - \Pr(A_1|T_1A_2T_2) .$$

This equation specifies the largest possible difference between first-order sequential predictions, and it is interesting to examine the effect of the information variable on the value of Δ. To further simplify our analysis we shall let $\gamma = 1/2$ and therefore $p_\infty = 1/2$ for both the information and no-information cases.

Using Eqs. 8a and 8d, we may derive an expression for Δ in the information case given that $p_\infty = \gamma = 1/2$; namely

$$\Delta(I) = \frac{(1-h)(1-h+2h\theta)}{N(1+h)} . \quad (12)$$

Similarly, using Eqs. 5a and 5c (with $\pi_1 = \gamma$, $\pi_2 = 1-\gamma$, and $\pi_0 = 0$) we may derive the following expression for Δ in the no-information case given, again, that $p_\infty = \gamma = 1/2$:

$$\Delta(\bar{I}) = \frac{(1-h)(1-h+c+2c'h)}{N(1+h)} \quad (13)$$

Comparing Eqs. 12 and 13 makes it apparent that the maximum possible difference between first-order sequential predictions may be greater or less in the information case, as compared with the no-information case, depending on the values of θ, c', and h. Namely

$$\Delta(I) > \Delta(\bar{I}) , \qquad \text{if} \quad \theta > c' + \frac{c}{2h}$$

$$\Delta(I) < \Delta(\bar{I}) , \qquad \text{if} \quad \theta < c' + \frac{c}{2h} .$$

Thus, at least for this particular comparison, no parameter-free conclusions can be drawn about the overall sequential effects in the information and no-information situations.

It is clear that more comparisons can be made between the feedback and no-feedback situation. However, for our present purposes the results already developed are sufficient to indicate the types of analyses that are possible in terms of the theory presented in this paper.

References

1. Blackwell, H. R. Psychophysical thresholds: Experimental studies of methods of measurement. Bull. Eng. Res. Inst., No. 36, Univ. of Michigan, 1953.

2. Swets, J.A., Tanner, W.P., and Birdsall, T.G. Decision processes in perception. Psychol. Rev., 1961, 68, 301-340.

3. Fechner, G. T. Elemente der psychophysik. Leip., Breitkopf and Hartel, 1860.

4. Carterette, E. C. and Wyman, M. J. Applications of a Markov learning model to a simple detection situation involving social pressure. In J. Criswell, H. Solomon and P. Suppes (Eds.), Mathematical Methods in Small Group Processes. Stanford, California: Stanford Univ. Press, 1962.

5. Atkinson, R. C. Mathematical models in research on perception and learning. In M. Marx (Ed.), Psychological Theory (2nd Edition). New York: Macmillan Co., 1962.

6. Atkinson, R. C. A variable sensitivity theory of signal detection. Technical Report No. 47, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1962.

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

8. Atkinson, R. C. and Estes, W. K. Stimulus sampling theory. In R. R. Bush, E. Galanter, and R. D. Luce (Eds.), Handbook of mathematical psychology. New York: Wiley, 1962, in press.

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

10. Atkinson, R. C. and Carterette, E. C. Signal detection as a function of the stimulus presentation schedule: a comparison of forced-choice and yes-no procedures (in preparation).