Implications of Latency Data for Threshold and Nonthreshold Models of Signal Detection[1]

EWART A. C. THOMAS

University of Michigan[2]

AND

JEROME L. MYERS

University of Massachusetts

Two detection cum latency models are constructed by combining sets of plausible assumptions about latency with, in the first instance, the model of the Theory of Signal Detection and, in the second instance, a general threshold model. The aim of this paper is to show that latency statistics can be used to provide sharp tests for distinguishing between the two models. The two statistics studied are the reaction time operating characteristic (RT-ROC) and the plot of average reaction time against response probability (RT-probability curve). It is shown that, under the first detection model, the RT-ROC lies below the ROC derived from detection rates except at the "yes–no" point, and the RT-probability curve is decreasing whereas, under the second model, the RT-ROC coincides with portions of the latter ROC, and decreasing RT-probability curves do not arise naturally.

The effect of criterion variability on the ROC is discussed in detail and conditions are given under which the RT-ROC is equivalent to a ROC generated by variable criteria.

Theories of signal detection have been almost exclusively concerned with choice proportions and ratings and, with rare exceptions, only those data have been collected. Only a very few investigators have publicly taken cognizance of the fact that responses require time for their execution. Much of that limited body of work shares the view that detection response times (RT) can be incorporated within the framework of the Theory of Signal Detection (Swets, Tanner, and Birdsall, 1961) or of the closely

[2] Now at Stanford University.

253

related strength theory of Wickelgren (1968); the assumption has generally been that RT decreases monotonically with distance from the cut-off (Bindra, Williams, and Wise, 1965; Bindra, Donderi, and Nishato, 1968; Norman and Wickelgren, 1969; Smith, 1968; Wickelgren, 1968). Other models have been considered by Carterette Friedman, and Cosmides (1965), Sekuler (1965), and Nickerson (1969). Only Nickerson, working with an extension of McGill's counter model (1963), has derived a detailed set of predictions. Consequently, we have limited knowledge of the implications of either the Theory of Signal Detection or threshold models (e.g., Luce, 1963; Krantz, 1969) for RT data.

On the other hand, for those models in which the decision variable, which determines what response is made, is regarded as the outcome of a temporal process (e.g., Carterette, 1966; Green and Luce, 1967; Laming, 1969), choice proportions and RT are related naturally and the implications for both sets of data can be derived, in principle.

In this paper, we will consider two classes of detection models and derive results for RT measures obtainable in "yes–no" experiments. One class consists of those models which view sensory experience as continuous. Within this class, most of our conclusions will apply to both the detection theory of Swets, Tanner, and Birdsall (1961), in which the decision variable is viewed as a likelihood ratio, and to the strength theory of Wickelgren (1968), in which the decision variable is viewed as a psychological construct. We will also show that our results hold under assumptions about the distribution of the decision variable which are more general than the usual normality assumption. The second class is of those models which assume that sensory experience consists of a set of orderable discrete states. Luce's two-state low threshold theory (1963) and Krantz's three-state low-high threshold theory (1969) are two examples of the general class of models with which we will be concerned.

It is clear from Krantz's discussion (1969) that it will be generally quite difficult, if not impossible, to discriminate between continuous models and a threshold model having at least three states on the basis of response measures usually analyzed in "yes–no" experiments—Receiver Operating Characteristic curves based on choice proportions or on confidence ratings, or posterior probabilities of stimuli given ratings. This difficulty in deciding between the two classes of models has in large part motivated our theoretical investigation of latencies. We will show that under quite general conditions, Receiver Operating Characteristics based on RT data will differ for continuous and certain threshold models. We will consider alternative assumptions about the relationship between RT and the decision variable for both classes of models and, in each specific case, we will derive parameter-free predictions about average RT as a function of experimental manipulations typical of much research in detection. Before presenting our derivations, we will briefly review the two classes of models.

A Model of Continuous Sensory Experience

Assumptions

The following assumptions have much in common with the Theory of Signal Detection (Swets, Tanner, and Birdsall, 1961) and with unidimensional strength theory (Wickelgren, 1968):

(i)   The subject's sensory experience on each trial is represented by a continuous random variable X, which takes on values along a psychological dimension x. In almost all of what follows, the x-axis will be presumed to be the decision axis. The special case in which X is regarded as a likelihood ratio will be considered only briefly.

(ii)   The probability density function (p.d.f.) of X given noise (n) and signal plus noise (sn) will be denoted by $f_n(x)$ and $f_s(x)$, respectively. For much of this paper it will be assumed that

$$f_n(x) = f(x), \tag{1a}$$

$$f_s(x) = f(x - \mu), \qquad \mu \geqslant 0, \tag{1b}$$

although some consideration will be given to the unequal variance cases, where

$$f_s(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right), \qquad \sigma \neq 1 \tag{1c}$$

and where $f_s(x)$ is the convolution of $f_n(x)$ and the p.d.f. of a positive random variable.

(iii)   For much of this paper, the only relevant property of $f(x)$ will be whether or not $f(x)$ belongs to a class $\mathscr{L}$, of distributions, where $f \in \mathscr{L}$ if and only if $-d^2 \log f(x)/dx^2 \geqslant 0$ for all x. $\mathscr{L}$ contains the normal, logistic, and gamma distributions (the exponential is the limiting case where the second derivative of $\log f(x)$ is zero) but does not contain "high-tailed" distributions such as the Cauchy, lognormal, and Pareto.

(iv)   There exists a criterion C, such that in a yes–no experiment the subject responds *yes* if $X \geqslant C$ and *no* otherwise. For the most part, it will be assumed that C has a fixed value c, throughout the experimental session, although some consideration will be given to the effects of trial-to-trial variation in C.

Letting $R(x) = \int_x^\infty f(y)\, dy$, etc., the response probabilities in a yes–no experiment with c fixed are

$$P_H = R_s(c)$$

and

$$P_{FA} = R_n(c),$$

where H and FA refer to hits and false alarms, respectively; CR and M will refer to correct rejections and misses.

RECEIVER OPERATING CHARACTERISTICS (ROC)

Fixed Criterion (FC)

It is assumed here that $f_n(x)$ and $f_s(x)$ are given by Eqs. (1a) and (1b), respectively, and that c is fixed within an experimental condition. Then the fixed criterion ROC based on *yes–no* data (FC–ROC) is the plot of $R(c - \mu)$ against $R(c)$ as c is varied over experimental conditions with μ held constant. Since $\mu \geqslant 0$, $R(c - \mu) \geqslant R(c)$ so that the FC–ROC lies above the positive diagonal and extends from $(0, 0)$ to $(1, 1)$. The slope of the curve is $f(c - \mu)/f(c)$ and it is not hard to prove the following important result.

THEOREM 1 (Lehmann, 1959, p. 330). *The slope of the* ROC *curve is monotonically decreasing if and only if* $f \in \mathscr{L}$.

If X, defined in assumption (i) above, is a likelihood ratio, then c, the likelihood ratio at the criterion, equals the slope of the ROC at the operating point generated by c. Therefore, the ROC slope is monotonic in c, so that $f(x)$ is automatically in $\mathscr{L}$.

We now consider the case where the signal distribution is the convolution of the noise distribution with the distribution of a *positive* random variable. In this case,

$$f_s(x) = \int_0^\infty f(x - y) g(y) \, dy, \qquad \text{where} \quad \int_0^\infty g(y) \, dy = 1.$$

The ROC slope when the criterion is c is

$$\frac{f_s(c)}{f_n(c)} = \int_0^\infty \frac{f(c - y)}{f(c)} g(y) \, dy.$$

Since $y \geqslant 0$, the integrand is monotonic in c iff $f \in \mathscr{L}$ (Theorem 1). Therefore, the ROC slope is monotonic if $f \in \mathscr{L}$. (This condition is sufficient but not necessary in this case.) It should be noted that in the normal-unequal variance model the signal distribution is the convolution of the noise distribution and another *normal* distribution, so that the condition for the above result is not satisfied. In fact, it is well known that for this model the ROC slope is nonmonotonic.

Variable Criterion

Consider an operating point (P_{FA}, P_H), generated by a criterion C which varies over trials; thus, for example,

$$P_{FA} = \int_{-\infty}^{\infty} e(c) \int_{c}^{\infty} f(x)\, dx\, dc = E[R(C)],$$

where $e(c)$ is a nondegenerate probability density function with finite mean and variance and "E" denotes "mathematical expectation." Assuming that both X and C are normally distributed, Wickelgren (1968) concludes that (a) the operating point will lie below the ROC generated by fixed criteria (FC–ROC) and (b) the degree of such depression will directly reflect the variance in the criterion generating the operating point. In one respect, Wickelgren's conclusions can be generalized. We will show that normality in the distribution of neither X nor C is necessary in order for the operating point to be depressed. The operating point will lie at or below the FC–ROC if (i) $f_s(x) = f(x - \mu)$, (ii) $f \in \mathscr{L}$, and (iii) the cut-off generating the operating point has a nondegenerate p.d.f.

In another respect, Wickelgren's conclusions must be qualified. In his discussion of criterion variability, he states that if ". . . the different methods do not differentially affect the strength distributions, then a different d' value tells us the lower operating characteristic has a greater criterion variance . . ." (1968, p. 110). The notion that increased variability causes decreased sensitivity is intuitively appealing and, in fact, will hold under many conditions. Nevertheless, it is important to realize that the conclusion is not necessarily true. We will show that there is at least one f in $\mathscr{L}$ for which depression of operating points is not monotonically related to the variability of criteria generating those points. Furthermore, even if X is normally distributed, it does not necessarily follow that the amount of depression at a given point on the ROC is monotonically related to the amount of criterion variability generating that point.

We first consider the general question of the effect of criterion variability upon the position of the ROC. It is not difficult to assess the result of variability if we make the simplifying assumption that the criterion randomly takes on two values over trials. Let $E_1, E_2, \ldots$ denote points on the ROC generated with *no* criterion variability and corresponding to criterion values of $c_1, c_2, \ldots$, respectively. Now consider a single experimental condition such that the criterion varies between c_1 and c_2 over trials. Then the operating point (P_{FA}, P_H) obtained under this condition will lie on the straight line $E_1 E_2$. Therefore, this point will always lie at or below the ROC joining E_1 and E_2 if and only if the ROC is concave, that is, if and only if $f \in \mathscr{L}$. The situation is depicted in the leftmost panel of Fig. 1.

The above heuristic argument can be generalized to the case where the criterion C has a nondegenerate p.d.f., $e(c)$, with finite mean and variance and $f_s(x) = f(x - \mu)$.
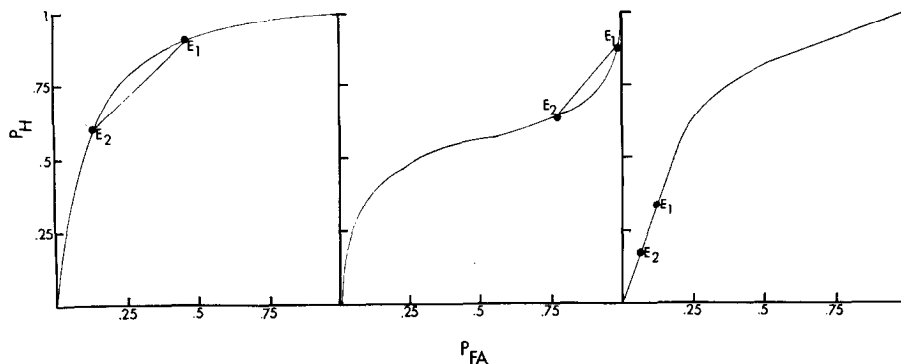
FIG. 1.   ROCs of different curvature, each showing an effect of criterion variability
on operating points.

We now prove the following theorem, which is a special case of a result on convex
functions:

THEOREM 2.   *For a given sensory model* $[f(x), f(x - \mu)]$, *every variable criterion
ROC lies uniformly at or below the* FC-ROC *iff* $-\log f(x)$ *is convex.*

Proof.   For ease of notation, let us make the substitutions

$$u = R(c), \qquad U = R(C),$$
$$v = R(c - \mu), \qquad V = R(C - \mu).$$

Then the FC-ROC is the plot of v against u and can be written as

$$v = L(u),$$

and, by Theorem 1, $L''(u) \leqslant 0$ iff $-\log f$ is convex.

Any point on a variable criterion ROC can be denoted by $[E(U), E(V)]$, where
the expectation is with respect to the p.d.f. of C. If $L''(u) \leqslant 0$,

$$L(u) \leqslant L'(u^*)(u - u^*) + L(u^*)$$

for fixed u^*, so that

$$E(V) = E[L(U)] \leqslant E[L'(u^*)(U - u^*) + L(u^*)].$$

When $u^* = E(U)$, this reduces to

$$E(V) \leqslant L[E(U)],$$

which, since $\{E(U), L[E(U)]\}$ lies on the FC-ROC, implies that the variable criterion
ROC lies at or below the FC-ROC.

If $L(u)$ is not concave, there exist two operating points $[u_1, L(u_1)]$ and $[u_2, L(u_2)]$ such that the straight line joining them lies above the FC–ROC. Choose $e(c)$ to be the discrete distribution with its mass located at $c = R^{-1}(u_1)$ and $c = R^{-1}(u_2)$. Then the operating point generated under this choice of $e(c)$ lies above the FC–ROC. This completes the proof.   Q.E.D.

The remaining two panels of Fig. 1 depict cases in which the degree of depression of the operating point generated by a varying criterion is not monotonically related to the degree of variability. The center panel depicts an ROC obtained when X is distributed normally; however, the signal distribution has greater variance. Because the first derivative does not decrease monotonically with P_{FA}, it is possible to obtain operating points generated by variable criteria which lie *above* the FC–ROC. In the rightmost panel, we have an FC–ROC derived from

$$f_n(x) = \tfrac{1}{2}e^{-|x|} \quad \text{and} \quad f_s(x) = \tfrac{1}{2}e^{-|x-\mu|}, \quad -\infty < x < \infty.$$

Then, for $0 < \mu \leqslant c$ and for $c \leqslant 0$ the ROC is linear with slopes e^{μ}, and $e^{-\mu}$, respectively; for $0 \leqslant c \leqslant \mu$ it is concave. Therefore, criterion variability will produce no depression in the ROC as long as $C \leqslant 0$ or $C \geqslant \mu$, but there will be a depression if $0 \leqslant C \leqslant \mu$. In other words, greater depression for central criteria does not necessarily imply that these criteria are more variable than other criteria.

In summary, if $f_s(x) = f(x - \mu)$, operating points generated by variable criteria lie uniformly at or below the FC–ROC iff $f \in \mathscr{L}$. Furthermore, the degree of depression depends not only upon the degree of criterion variability but also upon the curvature of the FC–ROC.

AN r-STATE MODEL

Several threshold models appear in the signal detection literature. Luce (1963) has suggested that r, the number of sensory states, equals 2; the stimulus either surpasses a single threshold or does not. To this sensory model Luce has added a decision mechanism. There are two possible states of bias: (i) *negative bias*, in which some detections result in *no* responses and (ii) *positive bias*, in which some failures to detect result in *yes* responses. Certain variables (e.g., signal probability, payoffs) determine the type and degree of bias, and varying these generates an ROC consisting of two linear segments. Krantz has proposed a model with three sensory states: nondetections, weak detections, and strong detections. He also assumes two types of bias (i) negative bias, in which some weak detections are converted into *no* responses and (ii) positive bias, in which some nondetections are converted into *yes* responses. The resulting ROC curve again consists of two limbs.

Still a third model might be considered, a three-state model with three states of bias, if we permit some strong detections to lead occasionally to *no* responses.

Then the resulting ROC would have three limbs. All three models that we have mentioned may be considered as a subset of a very general class of threshold models in which there are r states of sensory activation, including a nondetection state (cf. the neural quantum model of Norman (1964)), and r' states of bias, $r' \leqslant r$. In this section we will present a general formulation of the r-state model. In later sections, we will develop implications for response times.

We denote the activation states by $A_0, ..., A_{r-1}$, and let $q_i = P(A_i \mid n)$, $p_i = P(A_i \mid sn)$, $i = 0, ..., r - 1$. The A_i are thought of as being ordered on a sensory dimension such that the subject's sensory impression is greatest when he is in A_{r-1}, and he has no sensory information when he is in A_0. Thus it is reasonable to suppose that the likelihood ratio for sn over n given A_i increases with i, i.e., $p_0/q_0 < \cdots < p_{r-1}/q_{r-1}$. It is also assumed that there are r' "types" of bias $B_0, ..., B_{r'-1}$ such that for a given set of bias conditions, e.g., payoff and signal probability, the subject is in exactly one B_j. The response probabilities are then given by

$$P(\text{yes} \mid A_i, B_j) = \begin{aligned} &= 0 \quad & i < j \\ &= z \quad & i = j. \\ &= 1 \quad & i > j \end{aligned} \tag{2}$$

The bias state in an experimental condition is then characterized by the pair (B_j, z). We adopt the convention that B_j does not exist if $P(\text{yes} \mid A_j) = 1$ or 0 over all experimental conditions.

Given Eq. (2), we can readily derive expressions for hit and false alarm probabilities, conditional on the bias state (B_j, z):

$$P(H \mid B_j, z) = P_H^{(j)}(z) = p_j z + \sum_{i=j+1}^{r-1} p_i, \tag{3a}$$

$$P(FA \mid B_j, z) = P_{FA}^{(j)}(z) = q_j z + \sum_{i=j+1}^{r-1} q_i. \tag{3b}$$

As z varies for fixed B_j, we have a linear segment of the yes–no (YN) ROC, given by [cf. Norman, 1964, Eq. (6)]

$$P_H^{(j)} = \frac{p_j}{q_j} P_{FA}^{(j)} + \sum_{i=j+1}^{r-1} p_i - \frac{p_j}{q_j} \sum_{i=j+1}^{r-1} q_i. \tag{4}$$

Note that the YN–ROC consists of as many limbs as there are "types" of bias states. Thus the number of such limbs will be less than or equal to r. Krantz's model is an example in which the number of limbs are less than r because z always equals one when A_{r-1} is activated.

The variable criterion case discussed under the continuous model has an analog

under threshold models. An operating point will lie below the YN–ROC if the B_j generating the point varies and the likelihood ratio p_i/q_i increases monotonically with i. The condition on the likelihood ratio is equivalent to the condition that $f \in \mathscr{L}$ in the continuous case and serves the same function of guaranteeing a nonincreasing slope for the ROC curve. The conditions under which confidence ratings ROCs lie below YN–ROCs, discussed by Broadbent (1966) and Krantz (1969), are equivalent to assuming variable B_j and the general method of proof presented by Broadbent is the same as that used to demonstrate that variable B_j ROCs lie at or below YN–ROCs for any threshold model in which the likelihood ratio condition is met.

RT-ROC Curves

Constructing the RT–ROC

Several investigators (Norman and Wickelgren, 1969; Moss, Myers and Fillmore, 1970; Yager and Duncan, 1971) have constructed ROCs from RT data for a yes–no experiment. The procedure involves separating RT for sn and n trials. These sets of RTs are then ordered from the fastest *yes* response to the fastest *no* response. In brief, the latencies are ordered as if they represented an underlying continuum from strong *yes* to strong *no*. Next, a series of cut-points are selected along the RT dimension. This results in a series of ordered categories for each of the two stimuli. Category 1 would contain the fastest *yes* response, e.g., 0–200 msec.; category m, say, would contain the slowest *yes* response, e.g., those longer than 2000 msec.; category $m + 1$ would contain the slowest *no* responses, and categories with higher indices would contain faster *no* responses. An RT–ROC curve is obtained by plotting the points (x_k, y_k) where x_k and y_k are the proportions of RTs in the first k categories given n and sn, respectively. Note that y_m is the proportion of all RTs that are *yes* RTs and is therefore equal to P_H; similarly, x_m equals P_{FA}. Thus, the RT–ROC and YN–ROC curves must intersect at one point at least.

Continuous Models

Criterion variability and the RT–ROC. In what follows we assume that the latency data are obtained under a single, fixed decision cutoff c. Within the context of the continuous model, part of the variation in RT may be attributed to variation in $X - c$, the distance of the sampled observation from the cutoff. In addition, for any fixed distance $x - c$, we assume that RT is a random variable with a non-degenerate distribution function $G_{|x-c|}(t)$ and p.d.f. $g_{|x-c|}(t)$. It is implied, therefore, that this conditional distribution of RT does not depend upon the nature of the stimulus, and, for different values of c (induced, for example, by variations in payoffs or signal probabilities) but equal values of $x - c$, $g_{|x-c|}(t)$ will be invariant.

The equation of the RT–ROC is given by

$$x_k = \int_c^\infty f(x)\, G_{x-c}(t_k)\, dx, \qquad (5a)$$

and

$$y_k = \int_c^\infty f(x - \mu)\, G_{x-c}(t_k)\, dx, \qquad (5b)$$

where t_k is the upper bound of the latencies in category k ($\leqslant m$, the Eqs. for $k > m$ being similar). The two first derivatives of G are

$$g_x(t) = \frac{\partial}{\partial t}\, G_x(t),$$

and   (6)

$$\gamma_x(t) = \frac{\partial}{\partial x}\, G_x(t).$$

On integrating Eqs. (5a) and (5b) by parts, and replacing $x - c$ by x, we have

$$x_k = \int_0^\infty \gamma_x(t_k)\, R(x + c)\, dx, \qquad (7a)$$

$$y_k = \int_0^\infty \gamma_x(t_k)\, R(x + c - \mu)\, dx. \qquad (7b)$$

It has been argued, e.g., by Wickelgren (1968) and Norman and Wickelgren (1969) that the RT–ROC lies below the FC–ROC as a logical consequence of the analogy between the RT–ROC and an ROC generated by variable criteria. We will review this argument in an attempt to show that the analogy is misleading unless certain conditions are placed on the latency distribution function.

In the confidence rating experiment the subject is asked to maintain two or more criteria simultaneously. When the ROCs obtained from this experiment are found to lie below the FC–ROC it is argued that the cause of this depression is criterion variability, which, in turn, is presumed to be a consequence of the demands of maintaining two or more criteria. We note that given X, there exists a nondegenerate distribution of ratings (response) if and only if the criteria are variable. In a similar manner with respect to the RT–ROC, the p.d.f. $g_{|x-c|}(t)$ induces a conditional distribution of observed latency (response) given X. However, while there are two or more latency criteria, these criteria (unlike rating criteria) are fixed by the experimenter and, therefore, cannot be variable. For this reason the analogy between the RT–ROC and variable criterion ROCs can be misleading.

The possibility remains, nevertheless, that there exist conditions under which the

RT–ROC is, in some sense, equivalent to a variable criterion ROC. In order to explore this possibility it is helpful to have a formal definition of a variable criterion ROC. It seems natural to have at least two conditions which a variable criterion model should satisfy. The first is the obvious one that two distinct points on the same variable criterion ROC are generated by taking expectations of two distinct criterion distributions under the same pair of sensory p.d.f.s. The second, consistent with the notion that bias and sensitivity are independent, is that the criterion distribution is independent of the signal and noise distributions. Because of this second restriction, the formal definition of a variable criterion model involves the consideration of two or more signal distributions.

Let us consider any two ROCs, R and R', generated under signals s and s', respectively, and noise n. Let (u_1, v_1) and (u_2, v_2) denote two distinct points on R, and let (u_1, v_1') denote the point on R' having the same abscissa value as (u_1, v_1).

DEFINITION. R and R' are variable criterion (VC)–ROCs iff there exist distinct p.d.f.s $e_1(x)$ and $e_2(x)$ such that

(i) $\quad u_i = \int_{-\infty}^{\infty} e_i(x) R_n(x)\, dx \qquad i = 1, 2$

(ii) $\quad v_i = \int_{-\infty}^{\infty} e_i(x) R_s(x)\, dx \qquad i = 1, 2,$

and

(iii) $\quad v_1' = \int_{-\infty}^{\infty} e_1(x) R_{s'}(x)\, dx.$

Now considering the RT–ROC, let (x_k, y_k) and (x_k, y_k') be points on two RT–ROCs generated under the same (fixed) value of c but different signals s and s', respectively. According to the above definition, to assert that these two OCs are VC–ROCs is to assert that there exists a p.d.f. $e(x)$ such that

$$x_k = \int_0^\infty \gamma_x(t_k) R_n(x + c)\, dx = \int_{-\infty}^\infty e(x) R_n(x)\, dx, \qquad (8a)$$

$$y_k = \int_0^\infty \gamma_x(t_k) R_s(x + c)\, dx = \int_{-\infty}^\infty e(x) R_s(x)\, dx, \qquad (8b)$$

and

$$y_k' = \int_0^\infty \gamma_x(t_k) R_{s'}(x + c) = \int_{-\infty}^\infty e(x) R_{s'}(x)\, dx. \qquad (8c)$$

Letting

$$R_{s'}(x) = \begin{cases} 1 & x \leqslant y \\ 0 & x > y \end{cases},$$

(8c) reduces to

$$\int_0^y \gamma_x(t_k) \, dx = \int_{-\infty}^y e(x) \, dx,$$

from which it follows that

$$e(x) = \begin{cases} \gamma_{x-c}(t_k) & x \geqslant c \\ 0 & x < c \end{cases}.$$

Therefore, $\gamma_{x-c}(t_k) \geqslant 0$ and $\lim_{x \to \infty} G_{x-c}(t_k) = 1$.

If $\gamma_{x-c}(t) \geqslant 0$ and $\lim_{x \to \infty} G_{x-c}(t) = 1$, then, clearly, the RT–ROC is a VC–ROC. This proves the following:

THEOREM 3. *The RT–ROC is a VC–ROC iff latency decreases stochastically to zero as* $| x - c |$ *increases.*

From Theorems 2 and 3 we derive the following:

THEOREM 4. *If* $G_0(t) = 0$, $\lim_{x \to \infty} G_{x-c}(t) = 1$ *and* $\gamma_{x-c}(t) \geqslant 0$, *the* RT–ROC *lies uniformly at or below the* FC–ROC *iff* $-\log f$ *is convex.*

The last part of our characterization of the RT–ROC is to link the condition $\gamma_x(t) \geqslant 0$ to a desirable property of the RT–ROC, viz., that it should lie above the diagonal $v = u$, the line which describes chance performance. This property is desirable for if the RT–ROC for a subject does not lie wholly above the diagonal but his FC–ROC does, this would suggest that latency is an inappropriate measure of sensitivity for that subject.

We prove the following:

THEOREM 5. *The* RT–ROCs *for all sensory models* $[f(x), f(x - \mu)]$ *and all criterion values c lie above the diagonal* $v = u$ *iff* $\gamma_x(t) \geqslant 0$.

Proof. Since $R(x + c - \mu) \geqslant R(x + c)$, using Eq. (8), $\gamma_x(t) \geqslant 0$ implies that

$$y_k \geqslant x_k .$$

Now suppose that $\gamma_x(t_k) < 0$ for $a \leqslant x \leqslant b$. We show that there exists a sensory model and a criterion value for which $y_k \leqslant x_k$. Choose

$$\delta < \frac{b - a}{2} \quad \text{and} \quad \epsilon < \frac{G_a{}'(t_k) - G_b{}'(t_k)}{1 + G_a{}'(t_k) - G_b{}'(t_k)} ,$$

where

$$a' = a + \delta \quad \text{and} \quad b' = b - \delta.$$

Let $\mu = b - a - \delta$, and

$$f(x + c) = (1 - \epsilon)/\delta, \qquad a \leqslant x \leqslant a',$$
$$= f(x + c), \qquad \text{otherwise.}$$

Then

$$f(x + c - \mu) = (1 - \epsilon)/\delta, \qquad b' \leqslant x \leqslant b,$$
$$= f(x + c - \mu), \quad \text{otherwise.}$$

From Eq. (8),

$$x_k \geqslant [(1 - \epsilon)/\delta]\, \delta G_{a'}(t_k) = (1 - \epsilon)\, G_{a'}(t_k),$$

and

$$y_k \leqslant [(1 - \epsilon)/\delta]\, \delta G_{b'}(t_k) + \epsilon = (1 - \epsilon)\, G_{b'}(t_k) + \epsilon.$$

Therefore $y_k - x_k \leqslant 0$, which proves the theorem.   Q.E.D.

With regard to the ROCs of a subject, Theorems 4 and 5 together assert that, assuming $-\log f$ is convex, if all his RT–ROCs lie above the diagonal then they lie below the FC–ROC. However, it is possible that, for a *given* sensory model, we have *both* $\gamma_x(t) < 0$ for some x and the RT–ROC lying above the diagonal. In such a case, RT–ROC is not equivalent to VC–ROC. Nevertheless, under this more general condition, we can still prove that RT–ROC lies at or below FC–ROC as long as $f \in \mathscr{L}$. In short, even if latency does not decrease stochastically to zero as $|x - c|$ increases, RT–ROC lies beneath FC–ROC. We will now prove this.

Relation of the RT–ROC *to the* FC–ROC.   In this section we will show that, for any nondegenerate distribution function $G_{x-c}(t)$, the RT–ROC lies at or below the FC–ROC if $f_n(x)$ and $f_s(x)$ are given by Eqs. (1a) and (1b), where $f \in \mathscr{L}$. Consider the point (x_k, y_k) on the RT–ROC, and let us choose a point on the FC–ROC such that $x_k = P_{FA}$, where

$$P_{FA} = \int_{c'}^{\infty} f(x)\, dx. \tag{9}$$

Note that if $P_{FA} = x_k$, $c' > c$ unless $k = m$ ($t_k = \infty$); in that case, $G_{x-c}(t_k) = 1$ for all $x - c$, and $c' = c$. This is merely a restatement of the fact that the RT and FC–ROC curves meet at the "yes–no" point.

We wish to prove that

$$P_H = \int_{c'}^{\infty} f(x - \mu)\, dx \geqslant y_k.$$

We approach the problem by assuming that $G_{x-c}(t_k)$ varies in discrete steps as $x - c$

increases; the values of x at which $G_{x-c}(t_k)$ changes value are $c_1, c_2, ..., c_i, ..., c_n$. Let the cutoff, previously referred to as c, be c_0. The constant value of $G_{x-c}(t_k)$ between c_i and c_{i+1} is designated G_i ; then x_k and y_k are given by

$$x_k = \sum_{i=0}^{n-1} G_i(t_k) \int_{c_i}^{c_{i+1}} f(x) \, dx + G_n(t_k) \int_{c_n}^{\infty} f(x) \, dx \qquad (10a)$$

and

$$y_k = \sum_{i=0}^{n-1} G_i(t_k) \int_{c_i-\mu}^{c_{i+1}-\mu} f(x) \, dx + G_n(t_k) \int_{c_n-\mu}^{\infty} f(x) \, dx. \qquad (10b)$$

There exists a set of values, c_i' $(i = 0, 1, ..., n)$ such that

$$\int_{c_i'}^{c_{i+1}'} f(x) \, dx = G_i(t_k) \int_{c_i}^{c_{i+1}} f(x) \, dx, \qquad i < n \qquad (11a)$$

and

$$\int_{c_n'}^{\infty} f(x) \, dx = G_n(t_k) \int_{c_n}^{\infty} f(x) \, dx. \qquad (11b)$$

Since $G_n(t_k) \leqslant 1$, from Eq. (11b) we have $c_n' \geqslant c_n$. Putting $i = n - 1$ in Eq. (11a) and adding to Eq. (11b), we get

$$\int_{c_{n-1}'}^{c_n'} f(x) \, dx + \int_{c_n'}^{\infty} f(x) \, dx = G_{n-1}(t_k) \int_{c_{n-1}}^{c_n} f(x) \, dx + G_n(t_k) \int_{c_n}^{\infty} f(x) \, dx.$$

Therefore,

$$\int_{c_{n-1}'}^{\infty} f(x) \, dx \leqslant \int_{c_{n-1}}^{\infty} f(x) \, dx,$$

that is,

$$c_{n-1}' \geqslant c_{n-1}.$$

Similarly, we can show that

$$c_i' \geqslant c_i \qquad \text{for all } i.$$

Note that on summing Eq. (11) over i, we get

$$\int_{c_0'}^{\infty} f(x) \, dx = x_k ;$$

from which, using (9), we observe that $c_0' \equiv c'$. Therefore,

$$P_H = \sum_{i=0}^{n-1} \int_{c_i'-\mu}^{c_{i+1}'-\mu} f(x)\, dx + \int_{c_n'-\mu}^{\infty} f(x)\, dx. \tag{12}$$

Now consider the quantity

$$Q = \frac{\int_{c_i-\mu}^{c_{i+1}'-\mu} f(x)\, dx}{G_i(t_k) \int_{c_i-\mu}^{c_{i+1}'-\mu} f(x)\, dx} = \frac{R(c_i'-\mu) - R(c_{i+1}'-\mu)}{G_i(t_k)[R(c_i-\mu) - R(c_{i+1}-\mu)]},$$

where Q is the ratio of the $(i+1)$-th term in Eq. (12) to the $(i+1)$-th term in Eq. (10b). In order to prove that $P_H \geqslant y_k$, it suffices to show that $Q \geqslant 1$ for all i. Accordingly, note that $Q = 1$ if $\mu = 0$. Thus, if $dQ/d\mu \geqslant 0$, $Q \geqslant 1$ for $\mu \geqslant 0$. Differentiating,

$$\frac{dQ}{d\mu} = [G_i(t_k)]^{-1} [R(c_i - \mu) - R(c_{i+1} - \mu)]^{-2}$$

$$\times \{[f(c_i' - \mu) - f(c_{i+1}' - \mu)][R(c_i - \mu) - R(c_{i+1} - \mu)]$$

$$- [f(c_i - \mu) - f(c_{i+1} - \mu)][R(c_i' - \mu) - R(c_{i+1}' - \mu)]\}$$

$$> 0 \quad \text{iff}$$

$$\frac{f(c_i' - \mu) - f(c_{i+1}' - \mu)}{R(c_i' - \mu) - R(c_{i+1}' - \mu)} \geqslant \frac{f(c_i - \mu) - f(c_{i+1} - \mu)}{R(c_i - \mu) - R(c_{i+1} - \mu)}. \tag{13}$$

It can be shown (Thomas, 1971) that

$$[f(x_1) - f(x_2)]/[R(x_1) - R(x_2)]$$

increases with x_1 and x_2. Therefore, since $c_i' \geqslant c_i$, Eq. (13) holds and $dQ/d\mu \geqslant 0$ if $f \in \mathscr{L}$.

We can prove in a similar manner that

$$\int_{c_n'-\mu}^{\infty} f(x)\, dx \geqslant G_n(t_k) \int_{c_n-\mu}^{\infty} f(x)\, dx.$$

Since each component of P_H (the ordinate of the *yes–no* point) is greater than or equal to the corresponding component of y_k, $P_H \geqslant y_k$ (the ordinate of the RT–ROC point) where the two ROC points have the same abscissa value.

If the k-th interval on the RT scale contains *no* responses (i.e., $k > m$), we redefine x_k and y_k :

$$x_k = P_{FA} + \int_{-\infty}^{c} f(x)[1 - G_{x-c}(t_k)] \, dx,$$

$$y_k = P_H + \int_{-\infty}^{c} f(x - \mu)[1 - G_{x-c}(t_k)] \, dx.$$

The proof then follows as before.

The proof has been carried out for an arbitrary discrete function $G_x(t)$. As n, the number of values of $G_x(t)$, approaches infinity and the distance between c_i and c_{i+1} approaches zero, $G_x(t)$ is, in the limit, continuous. This completes the proof of the following:

THEOREM 6.   *The RT–ROC lies at or below the FC–ROC if $f \in \mathscr{L}$.*

The proof suggests a corollary prediction. Since $Q - 1$ increases with μ, so does $P_H - y_k$; thus the continuous model makes the stated prediction that the depression of the RT–ROC relative to the FC–ROC should be more pronounced for higher values of μ. Data presented by Moss *et al.* (1970) and Norman and Wickelgren (1969) appear to be consistent with this prediction. In both cases, RT–ROC plotted on normal–normal paper exhibited more pronounced peaking at the yes–no point when μ was larger.

Norman and Wickelgren (1969) have commented on the apparent "peak" of the RT–ROC at the *yes–no* point. Such a cusp may be due to the fact that they plotted their RT–ROC on normal–normal paper, for we can prove the following:

THEOREM 7.   *The slopes of the RT–ROC and the FC–ROC are equal at the yes–no point.*

Proof.   On differentiating Eqs. (8a) and (8b), we have

$$\frac{dy_k}{dx_k} = \frac{dy_k/dc}{dx_k/dc} = \frac{\int_0^{\infty} \gamma_x(t_k) f(x + c - \mu) \, dx}{\int_0^{\infty} \gamma_x(t_k) f(x + c) \, dx} . \tag{14}$$

As $k \to m$, $t_k \to \infty$ and $\gamma_x(t_k) \to 0$ for fixed x, subject to $\int_0^{\infty} \gamma_x(t_k) \, dx = 1$. In other words, $\gamma_x(t_k)$ tends to the degenerate distribution having all its mass at $x = 0$, from which it follows that

$$\lim_{k \to m} \frac{dy_k}{dx_k} = \frac{f(c - \mu)}{f(c)} = \text{slope of the FC–ROC at the yes–no point.}   \text{Q.E.D.}$$

Therefore, the meeting of the RT–ROC and the FC–ROC is not the sharp peak which is observed when the points are plotted on normal–normal paper.

Threshold Models

For this class of models we assume that the distribution of times to make a particular response depends on:

(i) the activation state A_i, which generated the response, but not on the signal energy that produced A_i; in particular, it is assumed that response time is independent of the stimulus parameters $\{p_i\}$ and $\{q_i\}$;

(ii) the bias state (B_j, z); this dependence will not be explicit when we consider the RT–ROC, but will be when we consider mean response times.

(iii) whether the response is yes or no; however, this dependence can be omitted without losing any generality.

Let us introduce the following notation:

$$p_{ik}(k \leqslant m) = \text{probability that a } \textit{yes} \text{ response occurs in category } k,$$
$$\text{given state } A_i.$$

$$p'_{ik}(k > m) = \text{probability that a } \textit{no} \text{ response occurs in category } k,$$
$$\text{given state } A_i.$$

$$P_{ik} = \sum_{l=1}^{k} p_{il}, \qquad k \leqslant m,$$

$$P'_{ik} = \sum_{l=m+1}^{k} p'_{il}, \qquad k > m.$$

From these definitions it can be seen that under bias state (B_j, z)

$$P_{im} = 0, \qquad i < j,$$
$$P_{jm} = z,$$
$$P_{im} = 1, \qquad i > j.$$

THEOREM 8. *The* RT–ROC *lies at or below the* YN–ROC *if* p_i/q_i *increases with* i.

Proof. *Case* 1. $k \leqslant m$.

Let us assume that the subject is in bias state (B_j, z). Then the coordinates of the RT–ROC are

$$x_k^{(j)} = \sum_{i=j}^{r-1} q_i P_{ik}, \tag{15a}$$

$$y_k^{(j)} = \sum_{i=j}^{r-1} p_i P_{ik}. \tag{15b}$$

Let us choose a point on the YN–ROC such that $P_{FA}^{(l)}(z') = x_k^{(j)}$, for some l and z'. Since $k \leqslant m$, $P_{FA}^{(l)} \leqslant x_m^{(j)}$ ($= P_{FA}^{(j)}$), so that $l \geqslant j$, and from Eqs. (4) and (15), we obtain after some simplification,

$$P_H^{(l)}(z') - y_k^{(j)} = \frac{p_l}{q_l} \left\{ \sum_{i=j}^{l} (q_i P_{ik}) - \sum_{i=l+1}^{r-1} q_i \right\} + \sum_{i=l+1}^{r-1} p_i - \sum_{i=j}^{r-1} p_i P_{ik}$$

$$= p_l \left\{ \sum_{i=j}^{l} \left(\frac{q_i}{q_l} - \frac{p_i}{p_l} \right) P_{ik} + \sum_{i=l+1}^{r-1} \left(\frac{p_i}{p_l} - \frac{q_i}{q_l} \right) (1 - P_{ik}) \right\}. \tag{16}$$

Since p_i/q_i increases with i, all the terms on the right hand side of (16) are positive.

Case 2. $k > m$.

For bias state (B_j, z) the RT–ROC coordinates are

$$x_k^{(j)} = x_m^{(j)} + \sum_{i=0}^{j} q_i P_{ik}', \tag{17a}$$

$$y_k^{(j)} = y_m^{(j)} + \sum_{i=0}^{j} p_i P_{ik}'. \tag{17b}$$

Choosing a point on the YN–ROC such that $P_{FA}^{(l)}(z') = x_k$, we have $l \leqslant j$ and from Eqs. (2) and (4) we obtain after some simplification,

$$P_H^{(l)}(z') - y_k^{(j)} = p_l \left\{ \sum_{i=0}^{l} \left(\frac{q_i}{q_l} - \frac{p_i}{p_l} \right) P_{ik}' + \sum_{i=l+1}^{j-1} \left(\frac{p_i}{p_l} - \frac{q_i}{q_l} \right) (1 - P_{ik}') \right.$$

$$\left. + \left(\frac{p_j}{p_l} - \frac{q_j}{q_l} \right) (1 - z - P_{jk}') \right\}. \tag{18}$$

Since $P_{jk}' < 1 - z$ the right-hand side of (18) is positive. This proves that the RT–ROC lies at or below the YN–ROC.   Q.E.D.

THEOREM 9.   (i) *If the subject is in state* (B_0, z) *the* RT–ROC *coincides with the* YN–ROC *above the yes–no point.*

(ii) *If the subject is in state* (B_{r-1}, z) *the* RT–ROC *coincides with the* YN–ROC *below the yes–no point.*

Proof. (i) When $j = 0$, $k > m$ implies $l = 0$. Putting $j = l = 0$ in (18),

$$P_H^{(0)}(z') - y_k^{(0)} = 0.$$

(ii)   When $j = r - 1$, $k \leqslant m$ implies $l = r - 1$. Putting $j = l = r - 1$ in (16),

$$P_H^{(r-1)}(z') - y_k^{(r-1)} = 0. \hspace{3cm} \text{Q.E.D.}$$

The implications of Theorems 8 and 9 may best be understood by considering three cases. First, we may have a model with a subthreshold state A_0 and a large number of states above A_0, each with very small or infinitesimal probability of activation. Under a liberal criterion such that a "no" response implies B_0, a straight line segment is generated between the points $(1 - q_0, 1 - p_0)$ and $(1, 1)$. With p_0 substantially greater than zero, prediction (i) applies.

Second, we may have a model with a subthreshold state A_0, and a small number of superthreshold states. Under experimental conditions such that the subject is in bias state B_{r-1}, a linear segment connects $(0, 0)$ to (q_{r-1}, p_{r-1}). If there is a moderate probability that noise alone can activate the highest threshold state, A_{r-1} (i.e., $q_{r-1} > 0$), prediction (ii), as well as (i), applies. Luce's (1963) low-threshold theory is a case in point.

Third, we may have a model in which the highest superthreshold state is not activated in the presence of noise alone $(q_{r-1} = 0)$. Prediction (i) still applies. However, it is doubtful that the subject would enter B_{r-1} since it needlessly lowers the hit rate when the false alarm rate is already zero. Thus, prediction (ii) would not apply. Krantz's three-state Low-High threshold model (1969) is a case in point. Note that if B_{r-1} does exist, the lowest limb of the ROC is vertical, coincident with the ordinate.

As Krantz has pointed out, it is unreasonable to conceive of a threshold model in which the signal never falls below the lowest threshold and, certainly, noise must frequently do so, i.e., $p_0 > 0$ and $q_0 > 0$. Since, in addition, it should be possible to establish conditions under which the subject adopts a liberal criterion, part (i) of Theorem 9 states an important prediction of threshold models. This prediction is quite opposed to that of continuous models which require the RT–ROC to lie strictly below the FC–ROC under the conditions stated earlier.

RESPONSE TIMES UNDER THE CONTINUOUS MODEL

Invariance of Response Probabilities

Consider a set of experimental conditions (combinations of signal-to-noise ratios and signal probabilities) such that P_H does not vary significantly over the set. Assume that the *sn* distributions are all of the same form and have the same variances across all conditions. Then they can all be mapped onto a single distribution $f(x)$, all having the same value of c. Under our assumption that the RT varies only as a function

of $x - c$, it follows that the distributions of RT_H (reaction times of hits) should not vary significantly over experimental conditions.

Suppose that the distributions have different variances: specifically, for one experimental condition

$$P_{FA} = \int_c^\infty f(x)\, dx$$

and, for a second condition,

$$P_H = \frac{1}{\sigma} \int_{c'}^\infty f\left(\frac{x - \mu}{\sigma}\right) dx, \qquad (\sigma \neq 1);$$

and $P_{FA} = P_H$. Then

$$P_{FA} = P_H = \int_{(c'-\mu)/\sigma}^\infty f(x)\, dx$$

and, therefore,

$$c = \frac{c' - \mu}{\sigma}.$$

Now, the mean hit RT is given by

$$P_H \overline{RT}_H = \int_{c'}^\infty \tau_{x-c'} f_s(x)\, dx, \qquad \text{where} \quad \tau_x = \int_0^\infty t g_x(t)\, dt$$

$$= \frac{1}{\sigma} \int_{c\sigma+\mu}^\infty \tau_{x-(c\sigma+\mu)} f\left(\frac{x - \mu}{\sigma}\right) dx$$

$$= \int_c^\infty \tau_{\sigma(y-c)} f(y)\, dy, \qquad \left(y = \frac{x - \mu}{\sigma}\right)$$

$$\lessgtr \int_c^\infty \tau_{y-c} f(y)\, dy = P_{FA} \overline{RT}_{FA}$$

if $\sigma \gtrless 1$ and τ_x is a monotone decreasing function of x. Thus, given two equal response probabilities, the one based on the distribution having the larger variance will be associated with the smaller mean RT if τ_x decreases as x increases. If variances are equal, the RTs should be equal.

The range of comparisons based on invariant probabilities may be further extended over both *yes* and *no* responses if we assume that $f(x)$ is symmetric. For example, if we have two equal response probabilities, P_{FA} and P_{CR} , possibly obtained under different experimental conditions, their associated RT distributions will be the same if we assume that $f_s(x)$ is given by (1b), $f(x)$ is symmetric, and the RT-distance function is symmetric about the cutoff, i.e., τ varies with the absolute value of $x - c$.

RT-Probability Curves

We next consider the plot of RT as a function of response probability which we will refer to as the RT-probability curve. If it is assumed that $f_s(x)$ is given by (1b), then the RT-probability curve for false alarm responses would be the same as that for hit responses. We will omit subscripts H and FA in the proof of the following theorem.

THEOREM 10.   *The RT-probability curve is decreasing if $f \in \mathscr{L}$ and if τ_x is a decreasing function of x.*

Proof.   Given c, the average false alarm RT is given by

$$RT_c = \int_c^\infty \frac{f(x)}{R(c)}\, \tau_{x-c}\, dx,$$

if the integral converges. For simplicity, we assume that τ_x is bounded. Then on integrating by parts,

$$RT_c = \tau_0 + \int_0^\infty \frac{d\tau_x}{dx} \frac{R(x+c)}{R(c)}\, dx. \tag{19}$$

Now

$$\frac{d}{dc} \frac{R(x+c)}{R(c)} = \frac{R(x+c)}{R(c)} \left[\frac{f(c)}{R(c)} - \frac{f(x+c)}{R(x+c)} \right] \leqslant 0 \qquad \text{for} \quad x \geqslant 0$$

if $f \in \mathscr{L}$ (Thomas, 1971, Theorem 2.4). Therefore, since $d\tau_x/dx \leqslant 0$,

$$\frac{d\overline{RT}_c}{dc} = \int_0^\infty \frac{d\tau_x}{dx} \frac{d}{dc} \left[\frac{R(x+c)}{R(c)} \right] dx \geqslant 0 \qquad \text{if} \quad f \in \mathscr{L}. \qquad \text{Q.E.D.}$$

If $f_s(x)$ is given by (1c) and σ varies with μ, inferences may still be drawn. Suppose that for two experimental conditions (denoted by primes), $P' > P$ and $\sigma' > \sigma$. Consider a third condition which is such that $P'' = P'$ and $\sigma'' = \sigma$. Then from the above theorem, $\overline{RT}'' < \overline{RT}$, and from the discussion of probability invariance $\overline{RT}' < \overline{RT}''$, from which it follows that $\overline{RT}' < \overline{RT}$. When P and σ relationships are in conflict, the situation is ambiguous.

Extension to Percentiles of the RT Distribution

Because of the nature of observed distributions of latencies, investigators frequently use the median, rather than the mean, as the basic measure, and results consistent with those derived for $\overline{RT}$ can be derived for median RT (or for any other percentile). If we assume that $G_{|x|}(t)$ decreases as $|x|$ increases, we can show that T_α, the RT

equal to or greater than 100α percent of the RTs, increases as response probability decreases or as the cutoff c increases. Let

$$H_c(t) = \int_c^\infty \frac{f(x)}{R(c)} G_{x-c}(t) \, dx.$$

By comparison with Eq. (19), it can be seen that if $f \in \mathscr{L}$,

$$\frac{d}{dc} H_c(t) \leqslant 0.$$

Therefore, the solution, T_α, of $\alpha = H_c(t)$, increases as c increases.

Interactions

We briefly consider here the effects of shifts in bias upon the quantities $\overline{RT}_{CR} - \overline{RT}_M$ and $\overline{RT}_H - \overline{RT}_{FA}$. This interaction of bias and latency measures is of interest because, as we will see later in this paper, a somewhat different prediction will hold for threshold models than for continuous models. To view the problem within the context of the continuous model, it is helpful to map both the *sn* and *n* distributions onto the same curve $f(x)$. If the two distributions have equal variances, we may conceive of a single distribution with two cutoffs, $x = c$ and $x = c - \mu$. Assume two RT functions, τ_{x-c} and $\tau_{x-(c-\mu)}$ differing only in their points of origin. As bias is varied (e.g., through the manipulation of signal probability of payoffs), c and $c - \mu$ are shifted by equal amounts. The question of whether, for example, $\overline{RT}_{CR} - \overline{RT}_M$ varies with this shift in cutoffs is really the question of whether $d^2\overline{RT}/dc^2 = 0$. From Eq. (19) it is clear that the second derivative will not be zero; there will be an interaction between the measure and the bias variable. We are unable to specify the nature of the interaction since it appears to depend upon $f(x)$, τ_x, and the placement of c.

We next consider a second possible interaction, that between signal-to-noise ratio and bias parameters. In this instance, we will assume that fixed bias implies a fixed likelihood ratio β. When RT is monotonically related to response probability, it will suffice to consider changes in response probability as μ varies for various values of *fixed* β.

THEOREM 11. *If $f_n(x)$ and $f_s(x)$ are given by* (1a) *and* (1b), *where $f \in \mathscr{L}$, and a subject maintains a constant likelihood ratio criterion, β, over changes in signal/noise ratio μ, and over changes in cutoff c, then*

$$\frac{dP_{FA}}{d\mu} \begin{array}{ll} \leqslant 0 & \text{if} \quad f(\mu) \leqslant f(0)/\beta \\ > 0 & \text{if} \quad f(\mu) > f(0)/\beta, \end{array}$$

where the x-scale is chosen so that the modal value of X is 0.

Proof. For simplicity, it is assumed that $f'(0) = 0$. Also note that $\mu \geqslant 0$ and, therefore, $f(\mu) \leqslant f(0)$.

$$\beta = f(c - \mu)/f(c), \tag{20}$$

and

$$dP_{FA}/d\mu = (dP_{FA}/dc)(dc/d\mu) = -f(c)(dc/d\mu).$$

Now, since β is fixed,

$$\frac{dc}{d\mu} = \frac{\partial\beta/\partial\mu}{\partial\beta/\partial c} = \frac{f(c)f'(c - \mu)}{f(c)f'(c - \mu) - f'(c)f(c - \mu)}, \tag{21}$$

where $f'(x) = (d/dx)f(x)$. We observe that since $f \in \mathscr{L}$, $(d/dx)[f'(x)/f(x)]$ is negative which implies that the denominator of the *RHS* of (21) is positive. Therefore, the sign of $dP_{FA}/d\mu$ is the same as that of $-f'(c - \mu)$. Since $f'(0) = 0$, $dc/d\mu \geqslant 0$ if and only if $c \leqslant \mu$ from Eq. (21). Since $f \in \mathscr{L}$, β is a monotonic function of c, so that $c \leqslant \mu$ implies $\beta \leqslant f(0)/f(\mu)$ [putting $c = \mu$ in Eq. (20)]. Therefore,

$$\frac{dP_{FA}}{d\mu} \leqslant 0 \qquad \text{if} \quad f(\mu) \leqslant f(0)/\beta \tag{22a}$$

$$> 0 \qquad \text{if} \quad f(\mu) > f(0)/\beta \tag{22b}$$

Q.E.D.

If $\beta < 1$, the condition in Eq. (22b) implies that $f(\mu) > f(0)$, which is impossible since $f(0)$ is the modal density; therefore, $dP_{FA}/d\mu \leqslant 0$ for all $\beta < 1$. For $\beta \geqslant 1$, $dP_{FA}/d\mu$ is positive for small μ and negative for large μ. When RT varies inversely with response probability, if the subject maintains a likelihood ratio criterion which is less than 1, RT_{FA} increases with increasing signal-to-noise ratio (increasing μ); if the criterion exceeds 1, the plot of RT_{FA} at first decreases and then increases as signal-to-noise ratio increases. The inequalities in Eqs. (22a) and (22b) are reversed for RT_{CR} (since $dP_{CR}/d\mu = -dP_{FA}/d\mu$). Using the preceding method of derivation we can also show that the plot of RT_H (RT_M) against signal-to-noise ratio is decreasing (increasing) for $\beta > 1$ and is increasing (decreasing) for $\beta < 1$.

Alternative Assumptions about RT

In the preceding discussion of RT-probability functions, we assumed that the expected RT decreases monotonically as a function of distance from the cutoff. Although this is the position generally held (Bindra *et al.*, 1968; Norman and Wickelgren, 1969; Smith, 1968), other assumptions are at least superficially plausible. One natural assumption is that τ decreases as a function of x. Then for fixed μ, as

c increases, slower RTs will be excluded from the computation of the expected value of "yes" RT; that is, RT_H and RT_{FA} will decrease as P_H and P_{FA} decrease. Mathematically,

$$\frac{d\overline{RT}_{FA}}{dc} = \frac{d}{dc}\left[\frac{1}{R(c)}\int_c^\infty f(x)\,\tau_x\,dx\right]$$
$$\leqslant 0$$

by Thomas (1971, Lemma 2.3) since τ_x is a decreasing function of x. The prediction is contradicted by results from Carterette, Friedman, and Cosmides (1965) and Sekuler (1965) and, indeed, by numerous observations of the effect of bias upon choice RT.

We could make the somewhat counterintuitive assumption that τ increases with increasing x. Then "yes" RTs would increase with c. However, "no" RTs would also increase in contradiction to available data; with increasing c, larger values of τ_x would be included in the average RT. This problem can be remedied by assuming that τ depends on x and the response, that τ increases with x for "yes" responses but decreases for "no" responses. Aside from the fact that such an assumption jars out intuition, there is still another problem. Under high signal-to-noise ratios (μ large), RT_H will be high since the response is based primarily on values of τ associated with large x. The prediction that more discriminable stimuli result in longer response times when responded to correctly is unpalatable. Thus, simple "strength" assumptions appear unreasonable. It is possible that some combination of "distance from cutoff" and "strength" assumptions might work; it is not clear, at this point, that such a departure from parsimony will be required of the data.

Response Times Under Threshold Models

Basic Equations

In this section, we take the r-state model, presented formerly, as our model of detection. We assume that the observed average reaction time is the sum of a decision time component ρ, which depends on A_i and (B_j, z), as stated earlier, and another component which includes both activation and motor times. It is assumed that this latter component is independent of A_i and (B_j, z) and it will be omitted from the analysis.

Thus, we may write

$$RT(\text{yes} \mid A_i, B_j, z) = \begin{cases} \sigma_{j,z} & i = j, \\ \rho_{i,j} & i \neq j. \end{cases} \tag{23}$$

If we assume that bias varies *continuously*, then the bias states $(B_j, 1)$ and $(B_{j-1}, 0)$

are equivalent and we can write $RT(\text{yes} \mid A_j, B_j, 1) = RT(\text{yes} \mid A_j, B_{j-1}, 0)$, i.e.,

$$\sigma_{j,1} = \rho_{j,j-1}. \tag{24}$$

The RTs for hits and correct rejections under B_j are given by

$$RT_H^{(j)}(z) = [P_H^{(j)}(z)]^{-1} \left\{ p_j z \sigma_{j,z} + \sum_{i=j+1}^{r-1} p_i \rho_{ij} \right\}, \tag{25}$$

$$RT_{CR}^{(j)}(z) = [P_{CR}^{(j)}]^{-1} \left\{ q_j(1-z)\,\sigma_{j,1-z} + \sum_{i=0}^{j-1} q_i \rho_{i,j} \right\}, \tag{26}$$

where

$$P_H^{(j)}(z) = p_j z + \sum_{i=j+1}^{r-1} p_i \tag{27}$$

and

$$P_{CR}^{(j)}(z) = q_j(1-z) + \sum_{i=0}^{j-1} q_i. \tag{28}$$

From (27) and (28) it can be seen that

$$P_H^{(j)}(z) - P_H^{(j+1)}(z') = p_j z + p_{j+1}(1-z') \geqslant 0,$$

and

$$P_{CR}^{(j+1)}(z') - P_{CR}^{(j)}(z) = q_j z + q_{j+1}(1-z') \geqslant 0. \tag{29}$$

Therefore, the sequences $\{P_H^{(j)}(z_j)\}$ and $\{P_{CR}^{(j)}(z_j)\}$, $0 \leqslant z_j \leqslant 1$ and $j = 0, 1,..., r-1$ are decreasing and increasing, respectively.

Invariance of Response Probabilities

We earlier remarked on a rather fundamental property of continuous models: If response probabilities obtained under different experimental conditions are equal, rather strong predictions about the associated mean RT follow. This is not generally true of the discrete-state models because there are many parameter-value-combinations which will yield a particular response probability but will yield different response times. For example, consider two sensory bias combinations, which give rise to the same hit probability. For simplicity, we work with Krantz's three-state model, and assume that both values of P_H lie on the lower limb of their respective yes–no curves. Then

$$RT_H^{(1)}(z) = \frac{1}{P_H^{(1)}(z)} \{ p_1 z \sigma_{1,z} + p_2 \rho_{2,1} \},$$

and

$$RT_H^{(1)}(z') = \frac{1}{P_H^{(1)}(z')}\{p_1'z'\sigma_{1,z'} + p_2\rho_{2,1}\}.$$

Setting $P_H^{(1)}(z) = P_H^{(1)}(z')$, that is,

$$p_1 z + p_2 = p_1'z' + p_2',$$

there appears little that can be said about the relationship between the two RTs.

RT-Probability Curves

When discussing the model of Signal Detection Theory, we found that, under the assumption that $f \in \mathscr{L}$ and that RT is inversely related to distance from the cut-off, the RT-probability curve is decreasing, consistent with studies by Carterette et al. (1965) and Sekuler (1965). We will show in this section that, under differing assumptions about ρ, decreasing RT-probability curves are not derived naturally from the r-state model.

Assumption I. $\rho_{i,j} \equiv \rho_i$ is independent of j. It follows from Eq. (24) that $\sigma_{j,1} = \rho_j$. From (25),

$$RT_H^{(j)}(1) = \left[\sum_{i=j}^{r-1} p_j\right]^{-1}\left[\sum_{i=j}^{r-1} p_i\rho_i\right]. \tag{30}$$

It is clear that a necessary and sufficient condition for $RT_H^{(j)}(1)$ to be an increasing function of j for all $\{p_i\}$ is

$$\rho_0 \leqslant \rho_1 \leqslant \cdots \leqslant \rho_{r-1}. \tag{31}$$

Putting $z = 0$ in Eq. (26),

$$RT_{CR}^{(j)}(0) = \left[\sum_{i=1}^{j} q_i\right]^{-1}\left[\sum_{i=0}^{j} q_i\rho_i\right].$$

Therefore, if the inequalities (31) are satisfied, $RT_{CR}^{(j)}(0)$ is also an increasing function of j. Similarly, it can be shown that if $RT_{CR}^{(j)}(0)$ is a decreasing function of j, then so is $RT_H^{(j)}(1)$. This proves the following:

THEOREM 12.   Under Assumption I, the latency-probability curve for hits and that for correct rejections cannot both be decreasing.

Assumption II. $\rho_{i,j} = \rho_{|i-j|}$, is a decreasing function of $|i - j|$. It follows that $\sigma_{j,1} = \rho_1$.

This assumption is analogous to the one, made when we were considering the continuous model, that latency is inversely related to distance from the cut-off. We will show that the latency-probability curve is not monotonically decreasing by using the following contradiction:

THEOREM 13.   *If the latency-probability curve is decreasing for all* $\{p_i\}$, *then*

(a) $\sigma_{j,z}$ *is a decreasing function of* z

(b) $\lim_{z \to 0} z\sigma_{j,z} = 0$.

THEOREM 14.   *If* $\sigma_{j,z}$ *is a decreasing function of* z *and if* $\lim_{z \to 0} z\sigma_{j,z} = 0$, *then either*

(a) $\sigma_{j,0} = \lim_{z \to 0} \sigma_{j,z}$ *is infinite,* $\lim_{z \to 0} d/dz(z\sigma_{j,z})$ *is infinite and* $\lim_{z \to 0} z^2\sigma_{j,z} = 0$,

or

(b) $\sigma_{j,0}$ *is finite and* $\lim_{z \to 0} z\sigma'_{j,z} = 0 = \lim_{z \to 0} z^2\sigma'_{j,z}$.

THEOREM 15.   *Whether* $\sigma_{j,0}$ *is finite or infinite,*

$$\lim_{z \to 0} \frac{dRT_H^{(j)}(z)}{dP_H^{(j)}(z)} \geqslant 0.$$

The conclusion of Theorem 15 contradicts the assumption of Theorem 13.

Proof of Theorem 13.   (a) Put $p_i = 0$, $i \geqslant j + 1$.
   Then from (25),

$$RT_H^{(j)}(z) = \sigma_{j,z} .$$

Therefore, $\sigma_{j,z}$ is a decreasing function.   Q.E.D.

(b)   Let $s_j = \lim_{z \to 0} z\sigma_{j,z}$. If the latency-probability curve is decreasing

$$RT_H^{(j)}(0) \leqslant RT_H^{(j+1)}(1).$$

$$\therefore \quad \sum_{i=1}^{r-1-j} p_{j+1}\rho_i + p_j s_j \leqslant \sum_{i=1}^{r-1-j} p_{j+1}\rho_{i-1} , \quad \text{where} \quad \rho_0 = \rho_1 . \quad (32)$$

$$\therefore \quad \sum_{i=1}^{r-1-j} \frac{p_{j+1}}{p_j} (\rho_{i-1} - \rho_i) \geqslant s_j .$$

If $s_j > 0$, we can choose $\{p_i\}$ such that $(p_{j+2}/p_j) \leqslant (s_j/\rho_1 - \rho_2)$ and $p_i = 0$, $i \geqslant j + 3$. Then (32) would be violated. Therefore,

$$s_j = 0. \quad (33)$$

Q.E.D.

Proof of Theorem 14.

LEMMA (l'Hôspital). *If* $\lim_{x\to 0} f(x)$ *and* $\lim_{x\to 0} g(x)$ *are both* 0 *or both* ∞ *and* $f'(x)$ *and* $g'(x)$ *never vanish for the same* x, *then*

$$\lim_{x\to 0}\frac{f(x)}{g(x)} = \lim_{x\to 0}\frac{f'(x)}{g'(x)}, \tag{34}$$

whenever the second limit exists.

    (a) *If* σ_0 *is infinite (omitting the subscript* j*), putting* $f(z) = \sigma_z$ *and* $g(z) = 1/z$ *in* (34) *and using* (33),

$$0 = \lim_{z\to 0} z\sigma_z = -\lim_{z\to 0} z^2\sigma_z' \qquad \left(\sigma_z' \equiv \frac{d}{dz}\sigma_z\right).$$

Next, putting $f(z) = z\sigma_z$ *and* $g(z) = z$ *in* (34),

$$\lim_{z\to 0}\frac{d}{dz}(z\sigma_z) \qquad \text{*is infinite if* } \sigma_0 \text{ *is infinite.*} \qquad \text{Q.E.D.}$$

    (b) *If* σ_0 *is finite, then*

$$\sigma_0 = \lim_{z\to 0}\frac{d}{dz}(z\sigma_z) = \sigma_0 + \lim_{z\to 0} z\sigma_z',$$

and, therefore, $\lim_{z\to 0} z\sigma_z' = 0 = \lim_{z\to 0} z^2\sigma_z'$.   Q.E.D.

Proof of Theorem 15.

$$\lim_{z\to 0}\frac{dRT_H^{(j)}(z)}{dP_H(z)} = \left[\sum_{i=j+1}^{r-1} p_i\right]^{-2} \left\{\lim_{z\to 0}\frac{d}{dz}(z\sigma_{j,z})\sum_{i=j+1}^{r-1} p_i + p_j \lim_{z\to 0} z^2\sigma_{j,z}' - \sum_{i=j+1}^{r-1-j} p_{j+1}\rho_i\right\}.$$

If σ_0 is infinite, from Theorem 14,

$$\lim_{z\to 0}\frac{dRT_H^{(j)}(z)}{dP_H^{(j)}(z)} = +\infty,$$

and, if σ_0 is finite,

$$\lim_{z\to 0}\frac{dRT_H^{(j)}(z)}{dP_H^{(j)}(z)} = \left[\sum_{i=1}^{r-1-j} p_{j+1}\right]^{-1} \left\{\sigma_{j,0} - \frac{\sum_{i=1}^{r-1-j} p_{j+1}\rho_i}{\sum_{i=1}^{r-1-j} p_{j+1}}\right\}. \tag{35}$$

Since $\sigma_{j,z}$ is decreasing,

$$\sigma_{j,0} > \sigma_{j,1} = \rho_1.$$

Therefore, since $\rho_1 > \rho_2 > \cdots > \rho_{r-1}$, by assumption, the *RHS* of (35) is positive.
                                          Q.E.D.

This completes the proof that the assumption that the latency-probability curve is monotonically decreasing leads to the contradictory conclusion that the curve is increasing when the bias states are "close to" $(B_j, 0)$ for all j.

The above results merit restatement. We examined the consequences of two alternative assumptions about ρ: (i) ρ depends on the activation state and not on B_j, and (ii) given A_i and B_j ($i \neq j$), ρ is a decreasing function of $|i - j|$. Under (i) it was shown that the latency-probability curves cannot be decreasing for both hits and correct rejections, and under (ii) it was shown that the curve for neither is decreasing. If it were possible to obtain experimentally points on the latency-probability curve that are close together, this result could then be used to distinguish the Threshold model from that of Signal Detection Theory. However, if the data points are not close together, it is possible that the *empirical* curves are decreasing, as is seen in the following theorem:

THEOREM 16. *Under Assumption* II, *the sequence* $\{RT_H^{(j)}(1)\}$ *is increasing if* $\{\sum_{i=j+k}^{r-1} p_i / \sum_{i=j}^{r-1} p_i\}$ *is a decreasing sequence in* j (k *fixed*).

Proof. Since $\rho_0 = \rho_1 \geqslant \rho_2 \geqslant \cdots \geqslant \rho_{r-1}$, we can write

$$\rho_j = \rho^* - \sum_{k=0}^{j} a_k, \qquad j = 0, 1, ..., r - 1,$$

where $a_k \geqslant 0$.

$$RT_H^{(j)}(1) = \frac{\sum_{m=j}^{r-1} p_m \rho_{m-j}}{\sum_{m=j}^{r-1} p_m}$$

$$= \rho^* - \frac{\sum_{m=j}^{r-1} p_m \sum_{k=0}^{m-j} a_k}{\sum_{m=j}^{r-1} p_m}.$$

Now

$$\sum_{m=j}^{r-1} p_m \sum_{k=0}^{m-j} a_k = \sum_{k=0}^{r-1-j} a_k \sum_{m=j+k}^{r-1} p_m$$

$$= \sum_{k=0}^{r-1-j} a_k R_{j+k}, \qquad \text{where} \quad R_{j+k} = \sum_{m=j+k}^{r-1} p_m.$$

If $\{R_{j+k}/R_j\}$ is a decreasing sequence in j (k fixed),

$$\frac{R_{j+k}}{R_j} \geqslant \frac{R_{j+k+1}}{R_{j+1}}.$$

Therefore, since $a_k \geqslant 0$,

$$\sum_{k=0}^{r-1-j} a_k \frac{R_{j+k}}{R_j} \geqslant \sum_{k=0}^{r-2-j} a_k \frac{R_{j+k+1}}{R_{j+1}}.$$

$\therefore$ From (36), $RT_H^{(j)}(1) \leqslant RT_H^{(j+1)}(1)$.   Q.E.D.

In other words, if data points are obtained under bias states "close to" $(B_j, 1)$, and if R_{j+k}/R_j decreases as j increases, then *the empirical latency-probability curve is decreasing*. The condition on $\{p_i\}$ is analogous to the earlier requirement that $f(x)$ have an increasing hazard rate, so that this result is not surprising.

TABLE I

Summary of Results for the Two Models

Statistic	Discrete-state model[a]	Continuous model[b]
RT-ROC	Coincides with the YN-ROC above the yes–no point if that point lies on the uppermost limb.	Lies below the FC-ROC, except at the yes–no point.
Equal Response Probabilities	Cannot generally predict relationships among associated RT.	RTs identically distributed. In the unequal variance case, faster mean RT is associated with larger variance.
Latency-Probability Curve	(i) If decision latency depends on A_i, but not on B_j, this curve cannot be decreasing for both hits and correct rejections.	(i) The curve is decreasing for all classes of RT (hits, etc.) if mean decision latency is a decreasing function of distance from the cutoff, $\mid x - c \mid$.
	(ii) If decision latency, given A_i, B_j, is a decreasing function of $\mid i - j \mid$, (a) the curve for neither hits nor correct rejections is decreasing, but (b) if data points are obtained under bias states "close to" $(B_j, 1)$, then the empirical curve is decreasing if $\sum_{i=j+k}^{r-1} p_i / \sum_{i=j}^{r-1} p_i$ is a decreasing function of j.	(ii) If decision latency is a monotonic function of x, the curve is decreasing for some classes of RT and increasing for others.
Interactions	If the yes–no points lie on the uppermost limb, $RT_M - RT_{CR}$ is invariant under changes in bias and/or signal strength.	$RT_M - RT_{CR}$, for example, is a function of both bias and signal strength.

[a] It is assumed that p_i/q_i increases with i.
[b] It is assumed, unless otherwise stated, that $f_s(x) = f_n(x - \mu)$, where $f_n \in \mathscr{L}$.

Interactions

The following discussion will be restricted to motivational conditions such that the bias state is (B_0, z); that is, the *yes–no* point lies on the highest limb of the YN–ROC. Under these conditions, a *no* response will occur only when A_0 is activated. The mean reaction times of both misses and correct rejections will be,

$$RT_M^{(0)}(z) = \sigma_{0,1-z} = RT_{CR}^{(0)}(z).$$

The equality of these two RT measures in this bias state is sharply opposed to the prediction generated by the continuous model—that the response (presumably correct rejections since $\mu > 0$) having the higher probability should be faster, and that the difference in RT should vary as bias varies. Note that even if we find grounds for assuming that activation and motor times differ for the two responses we still predict that the difference in RT is constant as z varies, provided we make the natural assumption that only the decision time component is affected by the degree of bias.

SUMMARY OF PREDICTIONS

We have covered much ground and it may help to summarize the main predictions generated by the two models for various aspects of RT data in Table I.

DISCUSSION

The preceding summary highlights one problem with discrete-state models; rather restrictive assumptions are required to obtain empirical latency-probability curves that are decreasing (see Theorem 16). It is difficult to go beyond this in evaluating the validity of the models. Very few experiments have been performed in which RT has been collected and analyzed in ways relevant to our predictions. Furthermore, these experiments have varied sufficiently in procedure to make it difficult to draw strong general conclusions. Carterette *et al.* (1965) worked with an auditory signal, Sekuler (1965) with a visual one, and Moss *et al.* (1970) with recognition memory for tones. In the first of these studies, the observers controlled the stimulus duration while the signal was available for only a brief duration in the other two studies. As noted earlier, both of the detection studies obtained decreasing latency-probability curves as signal probability was varied. Neither obtained an effect of signal strength upon RT which is surprising within the context of either model. However, it may be that the variation employed was slight relative

to the variability in response times. A more disturbing result was one obtained by Moss *et al.* As the interval between standard and comparison tone decreased (generally viewed as a manipulation of signal strength by investigators of recognition memory), all subjects showed consistent decreases in all four *RT* measures. The continuous model would require that hits and correct rejections, which increased in probability, become faster while false alarms and misses become slower. Under the discrete-state model, we would expect the *no* responses to be invariant.

It is hard to evaluate the results cited above. They rest on a handful of subjects and have not been replicated. There is a paucity of appropriate data on which to base an evaluation of the two models. We need two kinds of studies. One kind of study would involve a systematic and thorough manipulation of signal strength and probability, permitting us to look at RT–ROC corresponding to various limbs of the discrete model's YN–ROC, and permitting an analysis of latency-probability curves and interactions of the sort indicated in the preceding summary. It would also be helpful if we had conditions which could be expected to yield different variances of the decision variable within the context of the continuous model. This model predicts that for two responses of different probabilities, the advantage in latency for the more probable response should become smaller as the variance of its theoretical distribution declines with respect to that of the less probable response. A second kind of study would attack the latency predictions made for that case when response probabilities are equal. These predictions would seem to reflect a basic property of the class of continuous models considered in this paper. Presumably, some pilot work would be needed to establish the appropriate experimental conditions, but the effort appears to be worthwhile.

In closing, we would like to make clear that the predictions derived in this paper will not permit us to accept either type of model. Models that we have not considered —e.g., counter models such as those proposed by Laberge (1962) and McGill (1963),— may account for response times and probabilities as well as or better than the models that we have considered. The same may be said of models, within the classes considered, under somewhat different assumptions about the latency mechanisms. However, the classes of models that we have considered are broad, plausible, and consistent with much of the theoretical work in detection. Thus, the failure of our predictions would permit the rejection of models that are clearly not trivial.

REFERENCES

BINDRA, D., WILLIAMS, J. A., AND WISE, J. S. Judgments of sameness and difference: experiments on decision time. *Science*, 1965, 150, 1625–1627.

BINDRA, D., DONDERI, D. C., AND NISHATO, S. Decision latency of "same" and "different" judgments. *Perception and Psychophysics*, 1968, 3, 121–130.

BROADBENT, D. E. Two-state threshold model and rating scale experiments. *Journal of the Accoustical Society of America*, 1966, **40**, 244–245.

CARTERETTE, E. C. Random walk models for reaction times in signal detection and recognition. *Proceedings of the XVIIIth Congress of Psychology (Moscow)*, 1966, **16**, 84–95.

CARTERETTE, E. C., FRIEDMAN, M. P., AND COSMIDES, R. Reaction-time distributions in the detection of weak signals in noise. *Journal of the Acoustical Society of America*, 1965, **38**, 531–542.

GREEN, D. M., AND LUCE, R. D. Detection of auditory signals presented at random times. Perception and Psychophysics, 1967, **2**, 441–449.

KRANTZ, D. H. Threshold theories of signal detection. *Psychological Review*, 1969, **76**, 308–324.

LABERGE, D. A recruitment theory of simple behaviour. *Psychometrika*, 1962, **27**, 375–396.

LAMING, D. R. J. From choice-reaction times to psychophysics. 1969. Unpublished manuscript.

LEHMANN, E. L. *Testing statistical hypotheses*. New York: Wiley, 1959.

LUCE, R. D. A threshold theory for simple detection experiments. *Psychological Review*, 1963, **70**, 61–79.

McGILL, W. J. Stochastic latency mechanisms. In Luce, R. C., Bush, R. R., and Galanter, E., (Eds.), *Handbook of Mathematical Psychology*, New York: Wiley, 1963, pp. 309–360.

MOSS, S. M., MYERS, J. L., AND FILMORE, T. Short-term recognition memory of tones. *Perception and Psychophysics*, 1970, **7**, 369–373.

NICKERSON, R. S. 'Same'–'different' response times a model and a preliminary test. In Koster, W. G. (Ed), *Attention and Performance II*. Amsterdam: North-Holland Publishing Company, 1969. Pp. 257–275.

NORMAN, D. A. Sensory thresholds, response biases, and the neural quantum theory. *Journal of Mathematical Psychology*, 1964, **1**, 88–120.

NORMAN, D. A., AND WICKELGREN, W. Strength theory of decision rules and latency in retrieval from short-term memory. *Journal of Mathematical Psychology*, 1969, **6**, 192–208.

SEKULER, R. W. Signal detection, choice response times, and visual backward masking. *Canadian Journal of Psychology*, 1965, **19**, 118–132.

SMITH, P. T. Cost, discriminability and response bias. *British Journal of Mathematical and Statistical Psychology*, 1968, **21**, 35–60.

SWETS, J. A., TANNER, W. P., AND BIRDSALL, T. G. Decision processes in perception. *Psychological Review*, 1961, **68**, 301–340.

THOMAS, E. A. C. Sufficient conditions for monotone hazard rate—an application to latency-probability curves. *Journal of Mathematical Psychology*, 1971, **8**, 303–332.

WICKELGREN, W. Unidimensional strength theory and component analysis of noise in absolute and comparative judgments. *Journal of Mathematical Psychology*, 1968, **5**, 102–122.

YAGER, D., AND DUNCAN, I.-A. Signal detection analysis of luminance generalization in goldfish using latency as a graded response measure. *Perception and Psychophysics*, 1971, **9**, 353–355.