Psychological Review
1969, Vol. 76, No. 3, 308–324

THRESHOLD THEORIES OF SIGNAL DETECTION[1]

DAVID H. KRANTZ[2]

University of Michigan

Two-state low-threshold theory (2-LT), combined with a probabilistic response model, gives a plausible account of those features of confidence-rating detection data that have been cited as evidence against it. Nevertheless, there are data that conclusively reject 2-LT. A successful threshold theory must have both a low and a high threshold (and in consequence, must postulate at least three sensory states involved in detection). Three-state low- and high-threshold theory (3-LHT) is compatible with available detection data and is quantitatively quite similar to normal distribution signal detection theory with signal-plus-noise variance exceeding noise variance. The most promising way to reject 3-LHT, and all other threshold theories as well, would be to show that the posterior probability of the signal, given the response, can approach zero for moderate signal strengths.

Threshold theories characteristically assume that the presentation of a signal sometimes results in a state of nondetection (denoted $\bar{D}$). In State $\bar{D}$, the observer has no sensory basis for any sort of judgment about the signal. That is, given that the observer is in State $\bar{D}$, the probabilities of various responses are no different if the signal was present than if it was omitted.

The thresholds dealt with by threshold theories are internal, or observer thresholds—barriers separating internal sensory states. They are completely different from external, or energy thresholds. An observer threshold is not something to be measured in energy terms, it is something inferred from theoretical analysis of detection judgments. To show the logical independence of the two threshold concepts, suppose that there is an observer threshold. There may, nevertheless, be no energy threshold: for any nonzero signal energy, the probability of $\bar{D}$ may be less on signal than on nonsignal trials. On the other hand, even if there is no observer threshold, an energy threshold may exist: d' may be 0, for example, for some nonzero signal energies. This logical independence is illustrated in Figure 1, showing hypothetical psychometric functions for the four logical possibilities generated by the two-threshold–no-threshold dichotomies.

Here, we are concerned only with observer thresholds. Whether or not observer thresholds exist is an important question about information processing, akin to questions about all-or-none processes in memory, concept formation, etc. Of course, the answer to this question may conceivably vary with the modality, with absolute versus difference detection, with the degree to which observers are practiced, etc. At any rate, considerations having to do with energy thresholds are wholly beside the point.

This paper shall also distinguish sharply between sensory thresholds and response "thresholds" or criteria. It is presumed that response criteria or response biases are altered by motivation, instructions, or operant learning, while sensory thresholds are unaffected by these factors.

Evidence favoring threshold models was reviewed by Luce (1963a, 1963b). The main lines of evidence are (*a*) integer ratios between the abscissa value at which a psychometric function leaves 0.0 and that at which it reaches 1.0 (Stevens, Morgan, & Volkmann, 1941); and (b) chance recognition performance on nondetection trials in

[1] This paper was written while the author was a United States Public Health Service Special Postdoctoral Fellow at the Hebrew University in Jerusalem.

[2] The encouragement and helpful criticism of Amos Tversky, and valuable suggestions by David Green, Daniel Kahneman, R. Duncan Luce, Jacob Nachmias, Donald Norman, Irwin Pollack, Amnon Rapoport, John Swets, and Wayne Wickelgren are acknowledged. Requests for reprints should be sent to David H. Krantz, Department of Psychology, University of Michigan, Ann Arbor, Michigan 48104.

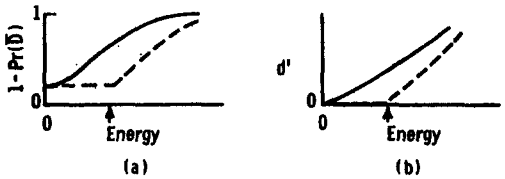

FIG. 1. Independence of observer and energy thresholds. (In both a and b, solid line indicates no energy threshold, dotted line indicates energy threshold at energy value shown by arrow. In a, there is an observer threshold, and $1 - Pr(\bar{D})$, the probability of exceeding threshold, is plotted. In b, there is no observer threshold and the standard signal detection measure d' is plotted.)

simultaneous detection-and-recognition experiments (Shipley, 1961). Luce proposed a two-state low-threshold model (2-LT). However, threshold models, and 2-LT in particular, have generally been rejected, on various grounds. This paper has four aims: to review the criticisms leveled at threshold theories and show that some of those criticisms are unjustified; to show that 2-LT is, nevertheless, incompatible with existing data; to present a new threshold theory; and to suggest a possible method for rejecting the existence of sensory thresholds, which does not depend on any special version of threshold theory.

CONFIDENCE RATINGS AND 2-LT

The principle arguments against 2-LT are based on data from confidence-rating (CR) detection experiments. As shall be shown subsequently, some CR data do reject 2-LT. However, the arguments that have been used heretofore are questionable. Since there has been some confusion on this point, and since other threshold theories should not be attacked by the same unjustified arguments, these arguments shall first be reviewed and an explanation given of how 2-LT gives a satisfactory account of many CR results.

Watson, Rilling, and Bourbon (1964) showed (for detection of a tone in white noise) that a 35-point CR Receiver Operating Characteristic (ROC) exhibited a curvilinear shape incompatible with the two-limbed rectilinear ROC predicted (for "yes-no" detection) by 2-LT. They state:

> The relation between the probability of correct detections and the probability of false alarms is described more closely by the ROC, as the latter is generated by overlapping normal distributions, than by certain other current models [p. 288].

The "other current models" are threshold models, particularly, 2-LT. Later, in response to criticism by Larkin (1965), Watson and Bourbon (1965) grant that the prediction of a two-limbed rectilinear ROC by 2-LT for CR data requires an additional assumption, but they still interpret their findings as evidence against two-state models.

The additional assumption, under which a two-limbed rectilinear ROC is predicted, is that the subject (*S*) divides the set of possible confidence responses into two subsets, using one subset exclusively in the nondetection state and the other subset exclusively in the detection state. Broadbent (1966) regards this assumption as the natural generalization of Luce's response model for "yes-no" detection. Moreover, Broadbent proved that any departure from this assumption predicts CR-ROC curves showing poorer performance than "yes-no" ROC curves, and he cites findings (e.g., Egan, Schulman, & Greenberg, 1959) contrary to this prediction.

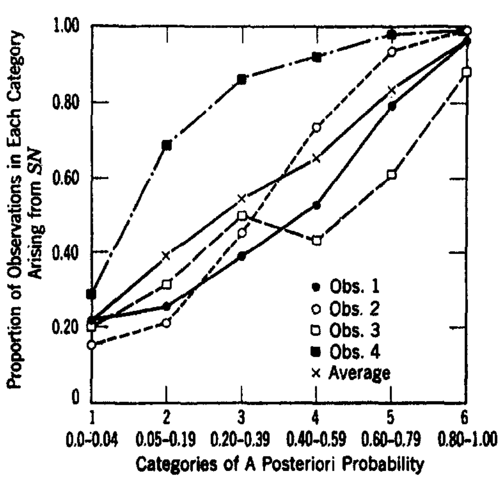

FIG. 2. Results of a CR visual detection experiment by Swets, Tanner, and Birdsall, 1961; reprinted with permission of the American Psychological Association. (The ordinate shows estimated posterior probability of signal, as a function of observer's confidence—expressed as one of six categories of subjective posterior probability.)

A slightly different type of analysis has also been interpreted as rejecting 2-LT. Swets (1961) cites the finding that the posterior probability of the signal, given a confidence response, is a monotone increasing function of confidence. The data he cites (from a visual detection experiment by Swets, Tanner, & Birdsall, 1961) are reproduced in Figure 2. Swets (1961) argues from these data: "The implication is that the human observer can distinguish at least six categories of sensory excitation [p. 172]." With respect to 2-LT, he concludes: "on the face of it, a two-category theory is inconsistent with the six categories of sensory excitation indicated by the rating data . . . [p. 174]." Similar arguments were advanced by Nachmias and Steinman (1963), from visual detection data, and by Norman and Wickelgren (1965), from recognition memory data. These authors, however, made explicit their use of supplementary assumptions about the way in which confidence responses are assigned to detection and nondetection states.

A third type of analysis was used by Sperling (1965), who showed (in a visual masking experiment) that higher confidence categories are used progressively more often as stimulus energy increases. Without giving a formal derivation, he erroneously concludes: "No 'threshold' theory of detection is adequate to account for the data [p. 555]."

In their recent book, Green and Swets (1966) repeat some of these arguments, and treat 2-LT as though it were simply inapplicable to CR detection.

In the next subsection, the predictions for CR data from 2-LT are systematically presented, and compared with the arguments cited above and with other features of existing CR data. Parts of the discussion were anticipated by Larkin (1965) and by Wickelgren (1968).

Analysis of CR Detection in the Framework of 2-LT

Prediction of detection performance requires the specification of both a sensory and a decision model. For classical signal detectability theory (SDT), the sensory model involves a continuum of sensory states (which will be denoted A) and two probability distributions over A: $P(a|SN)$ is the probability (or probability density) that State a (in A) occurs on a signal-plus-noise trial, while $P(a|N)$ is the probability (or probability density) that a occurs on a trial with noise alone. The two probability distributions are generally assumed to be Gaussian, with means μ_{SN} and μ_N and variances σ_{SN}^2 and σ_N^2. The response model for SDT assumes that A is partitioned into n (mutually exclusive and exhaustive) subsets, $A_1, \cdots, A_n$, and that S makes Response i if and only if the sensory state falls in Subset A_i. For "yes-no" detection, $n = 2$, while in CR experiments, the responses are $1, 2, \cdots, n$, where 1 will denote greatest confidence that the signal was present, 2, less confidence, $\cdots$, and n, least confidence (or greatest confidence that noise alone was present). Thus, SDT assumes a continuum of sensory states, and a deterministic response model (the response is determined strictly by the sensory state, once the partition, $A_1, \cdots, A_n$, has been set up).

For 2-LT, the sensory model involves two states, $\bar{D}$ (nondetection) and D (detection), plus two independent probabilities, $q = P(D|N)$ and $p = P(D|SN)$. Naturally, we assume $p > q$; the key sensory assumption differentiating 2-LT from classical high-threshold theory (HT) is that $q > 0$, rather than $q = 0$. (Figure 1a is a hypothetical plot of p against energy of signal; the intercept on the vertical axis is q.)

A deterministic response model for 2-LT would assign one response to D and another (or conceivably, the same) response to $\bar{D}$; thus no more than two confidence ratings would occur for a given response strategy. Since Ss use more than two confidence ratings in any given situation (they would probably feel quite silly if they did not), we are led to assume a probabilistic response model: $P(i|D) = \sigma_i$, $P(i|\bar{D}) = \tau_i$, where $\sigma_i, \ \tau_i \geq 0$ and $\sum_{i=1}^{n} \sigma_i = \sum_{i=1}^{n} \tau_i = 1$.

This response model is conveniently displayed as a matrix, whose rows sum to 1:

$$\begin{array}{c|cccc} & 1 & 2 & \cdots & n \\ \hline D & \sigma_1 & \sigma_2 & \cdots & \sigma_n \\ \bar{D} & \tau_1 & \tau_2 & \cdots & \tau_n. \end{array} \qquad [1]$$

Note that the choices of a continuous or discrete sensory model and a deterministic or probabilistic response model are logically separate. In fact, Lee (1963) explored the combination of continuous sensory model with probabilistic response model. However, the combination of two-state sensory and deterministic response models seems untenable.

Various special restrictions on Matrix 1 can be suggested. The one introduced above, proposed by Broadbent (1966) and Watson and Bourbon (1965), is that for each i, $\sigma_i = 0$ or $\tau_i = 0$. This proposal shall be referred to as the *nonoverlapping response model*, since it postulates that a response which occurs in State $D(\sigma_i > 0)$ cannot occur in State $\bar{D}$, and vice versa.

In deriving predictions for the "yes-no" case ($n = 2$), Luce (1963a) postulated that either $\sigma_1 = 1$ or $\tau_1 = 0$. This restriction can be justified in various ways. Intuitively, it is fairly plausible that S will not, on different trials of the same experiment, with constant instructions, payoffs, etc., both say "no" in $D(\sigma_1 < 1)$ and say "yes" in $\bar{D}(\tau_1 > 0)$. Also an S who is trying to maximize hit rate for a fixed false-alarm rate (Neyman-Pearson criterion) must satisfy exactly this restriction. Finally, direct experimental support was provided (using a two-state artificial sensory stage with a human decision stage) by Wickelgren (1967).

Other restrictions are possible. A natural generalization of Luce's "yes-no" restriction, to the case $n \geq 3$, is to postulate that the degree of overlap is at most $= 1$, that is, there is at most one response that occurs in both D and $\bar{D}$, with higher confidence responses occurring only in D and lower confidence ones only in $\bar{D}$. This postulate was used by Nachmias and Steinman (1963). A postulate that includes all the foregoing ones as special cases is that σ_i/τ_i is a nondecreasing function of confidence, that is, σ_i/τ_i is nonincreasing as i increases from 1 to n. The psychological justification of these latter postulates will be briefly considered below.

It shall be shown subsequently that the arguments against 2-LT, cited above, are insufficient to disprove two-state theory, if the nonoverlapping response model is replaced by the more general assumption that σ_i/τ_i is nonincreasing with i. Thus, it can be concluded, at the very least, that these arguments cannot be used to reject a sensory model without careful consideration of the response models. With this conclusion, few will disagree. However, many arguments have been proposed which favor the nonoverlapping response model. If these arguments are accepted, then the data cited above will stand at least as strong presumptive evidence against the two-state model. Therefore the arguments used in favor of the nonoverlapping response model shall be examined and it shall be shown that this model cannot be strongly relied upon.

One argument is that nonoverlapping response is the natural generalization of the deterministic response model of SDT. This argument scarcely needs refutation, since the "naturalness" of a model, in a formal sense, does not permit us to rely on the validity of that model. Still, it may be worth pointing out that the "naturalness" is somewhat illusory. For $n \geq 3$, the nonoverlapping response model is still probabilistic and not deterministic. It partitions responses, while SDT partitions sensory states; there is no logical connection between these two types of partition.

Second, nonoverlapping response has been taken to be the natural extension of Luce's "yes-no" model. This is incorrect, since the latter has overlapping responses. Moreover, the other justifications of the "yes-no" case are inapplicable when three or more responses are involved.

Third, it can be argued that some sort of sharp restriction is necessary to make two-state theory testable: It will not do to have $2(n - 1)$ free response parameters as in Matrix 1. This is not a strong argument for any particular sharp restriction. Moreover, as will be shown below, 2-LT is testable, and rejectable, with no restriction at all on Matrix 1. (See Equation 5 and the related discussion.) Finally, one can always reduce the problem of too many response parameters by using a number of different signal strengths interspersed randomly in a given experiment; this design requires the response model to remain constant over several signals and can reduce the ratio of free parameters to independent predictions.

Apart from refuting these particular arguments, it is well to consider briefly the psychological processes underlying confidence judgments. Consider an S who is in State D on a subset of trials, and has available several levels of confidence response. If the experimental situation implicitly demands that he distribute his responses, he will very likely comply. Wickelgren (1968) argues that such an S would be quite uncomfortable. The present author disagrees. The S can easily find a variety of "irrelevant" factors to serve as a basis for varying his confidence responses. These "irrelevant" factors (e.g., his momentary state of attention or alertness, etc.) may correspond to a multiplicity of internal states; the point is that the states are not *sensory* states, that is, they are not correlated with presence or absence of a signal. The S may not even be able to classify his internal states sharply into sensory and nonsensory components. If he cannot, it seems very likely that the response distribution in D will overlap the distribution in $\bar{D}$. Moreover, various other factors may promote overlap, including memory lapses (for previous response strategy) and imperfect discriminability of response (as in the mechanical analog to a rating scale used by Watson, Rilling, & Bourbon, 1964). Thus, the applicability of the nonoverlapping response model may vary sharply with the particular task.

Finally, it is worth noting that each of the more general restrictions on the Matrix 1 mentioned above is derivable from a simple psychological model of the response process. Suppose that S's internal states are of form (D,θ_j) or $(\bar{D},\theta_j)$, where D, $\bar{D}$ are detection and nondetection states and $\theta_1, \cdots, \theta_m$ are "attention" states, which are statistically independent of D, $\bar{D}$ and SN, N. If the total number of states, $2m$, exceeds the number of responses, n, then S has the possibility of responding deterministically, defining a response from the set $\{1, \cdots, n\}$ for each (D,θ_j) or $(\bar{D},\theta_j)$. Two plausible restrictions are the following: (*a*) The S has at least as much confidence in (D,θ_j) as in $(\bar{D},\theta_j)$, for each $j=1, \cdots, m$; (*b*) if θ_j is a higher attention state than θ_k, then S has at least as much confidence in the signal in (D,θ_j) as in (D,θ_k), and has at least as much confidence in *no signal* in $(\bar{D},\theta_j)$ as in $(\bar{D},\theta_k)$. From a and b it follows readily that there is at most one value of σ_i/τ_i between ∞ and 0; a plausible specialization is to assume that for some response i_0, $\tau_i = 0$ for $i < i_0$ and $\sigma_i = 0$ for $i > i_0$, while σ_i/τ_i can be between 0 and ∞ for $i = i_0$.

To derive the most general restriction, that σ_i/τ_i is a nonincreasing function of i, suppose that choice probabilities for responses $1, \cdots, n$ depend on response strength, $v(1), \cdots, v(n)$ according to Luce's model (1959) of multiple-choice behavior:

$$P(i) = v(i)/\sum_{j=1}^{n} v(j).$$

The response-strength matrix for D, $\bar{D}$ can be rewritten in the following form:

$$\begin{array}{c|cccc} & 1 & 2 & \cdots & n \\ \hline D & \alpha_1 v_1 & \alpha_2 v_2 & \cdots & \alpha_n v_n \\ \bar{D} & v_1 & v_2 & \cdots & v_n \end{array},$$

where $v(i|\bar{D}) = v_i$, $v(i|D) = \alpha_i v_i$. This model is conceptually similar to Luce's choice model for "yes-no" detection (1959, p. 60), except that SN, N in his response-strength matrix have been replaced by D, $\bar{D}$, respectively. It is highly plausible that the effect of moving from $\bar{D}$ to D would be to increase response strength more and more for high-confidence responses and decrease it more and more for low-confidence responses; that is, α_i should be a nonincreasing function of i. But

$$\sigma_i/\tau_i = \frac{\alpha_i v_i / \sum_{j=1}^{n} \alpha_j v_j}{v_i / \sum_{j=1}^{n} v_j} = \text{constant}\cdot\alpha_i,$$

so σ_i/τ_i is also a nonincreasing function of i.

We may conclude, then, that the nonoverlapping response model is not easily justifiable on formal or psychological grounds, and that competing, less restrictive hypotheses are at least quite plausible. Moreover, there is no need to assume the nonoverlapping response model to obtain sharp tests of 2-LT.

The ROC and the posterior probability functions for 2-LT shall be next derived, using the response model (Matrix 1).

The false-alarm rate, x_k, obtained by cumulating the first k responses as "yes" is

$$x_k = \sum_{i=1}^{k} P(i|N).$$

The corresponding hit rate, y_k, is defined similarly, replacing N by SN. Combining the sensory and response models, we have

$$\begin{aligned} P(i|N) &= q\sigma_i + (1-q)\tau_i \\ P(i|SN) &= p\sigma_i + (1-p)\tau_i. \end{aligned} \qquad [2]$$

Thus, the kth point on the ROC curve has coordinates (for $k = 1, \cdots, n-1$)

$$\begin{aligned} x_k &= q\sum_{i=1}^{k}\sigma_i + (1-q)\sum_{i=1}^{k}\tau_i, \\ y_k &= p\sum_{i=1}^{k}\sigma_i + (1-p)\sum_{i=1}^{k}\tau_i. \end{aligned} \qquad [3]$$

It is easily verified that three successive ROC points, (x_{k-1},y_{k-1}), (x_k,y_k), and (x_{k+1}, y_{k+1}) are collinear if and only if the ratios σ_k/τ_k and σ_{k+1}/τ_{k+1} are equal. Thus, the predicted CR-ROC curve will consist of two straight lines if and only if σ_i/τ_i takes on exactly two values, one value for $i \leq i_0$, another for $i > i_0$. The predicted curve coincides with the "yes-no" curve if and only if the only values of σ_i/τ_i are $+\infty$ and 0; otherwise (as Broadbent, 1966, pointed out), the CR-ROC points lie below the "yes-no" ROC curve.

The data of Watson, Rilling, and Bourbon (1964), which fail to conform to the two-straight-line prediction, also show CR performance clearly below the "yes-no" performance (Table I of their paper). Thus, these data are quite consistent with the possibility that 2-LT is valid and σ_i/τ_i takes on several values for a 35-point ROC curve. On the other hand, the data of Egan, Schulman, and Greenberg (1959) and of Nachmias (1968), which show equal performance in the "yes-no" and CR situation, are based on a four-category CR scale. The resulting three points for a CR-ROC cannot very well be used to reject 2-LT. (Markowitz & Swets, 1967, also compared four-category CR with "yes-no" ROC curves; their results, which are complex, are equivocal with respect to the present point.) At any rate, it seems reasonable, if 2-LT holds, that *S*s may have approximated $\tau_1 = \tau_2 = \sigma_3 = \sigma_4 = 0$ in the experiment by Egan et al. and by Nachmias, while responding with greater overlap on the finer scale of the Watson et al. study.

In any case, the lowered performance on the CR task in the latter experiment may be a slight embarrassment to SDT; it is perhaps best handled by a probabilistic response model, even with a continuum of sensory states (see Larkin, 1965, for examples). The detailed form of the Watson et al. curves also differs from that predicted by SDT, in precisely the way that would be expected if these curves were produced by "blurring" a two-limbed rectilinear curve: There is a definite suggestion of a "corner" in the region of the negative diagonal.

To derive the posterior probability of the signal, given i, we proceed as follows. Let $\Omega = P(SN)/P(N)$ be the (experimenter-controlled) prior odds of the signal, and let $\Omega_i = P(SN|i)/P(N|i)$ be the posterior odds of the signal, given Response i. By Bayes' theorem,

$$\Omega_i/\Omega = P(i|SN)/P(i|N).$$

Substituting Equation 2 in the above gives

$$\frac{\Omega_i}{\Omega} = \frac{p\sigma_i + (1-p)\tau_i}{q\sigma_i + (1-q)\tau_i}. \qquad [4]$$

From Equation 4, it is easily verified (assuming $p > q$) that Ω_i is a strictly increasing function of σ_i/τ_i. Thus, the posterior probability takes as many values as there are distinct values of σ_i/τ_i. If, as suggested above, σ_i/τ_i is monotonically decreasing with i, then posterior probability will be a monotonic increasing function of confidence that the signal was presented. Swets' argument, that multiple gradations of posterior probability are impossible with just two sensory states, is valid only if σ_i/τ_i is restricted to take exactly two values.

Similarly, the argument against two-state theory by Norman and Wickelgren (1965), depends heavily on the nonoverlapping response model. (The overstrong conclusions have since been retracted by Wickelgren, 1968.) Fine gradations of posterior probability as a function of confidence is simply not a strong enough datum to reject a two-state sensory model.

Nachmias and Steinman's analysis of CR performance is related to the instructions given their *S*s. In terms of the present numbering (which is reversed in their study), they told each *S* to respond on a 5-point scale, using Responses 1 or 2 when he saw something, 4 or 5 when he did not, and 3 when he was quite uncertain. They assumed that if 2-LT were true, *S* would interpret these instructions by setting $\tau_1 = \tau_2 = \sigma_4 = \sigma_5 = 0$. It would then follow (Equation 4) that $\Omega_1 = \Omega_2$ and $\Omega_4 = \Omega_5$, which is grossly false in their data.

On the other hand, it is far from certain that an *S* will conform without fail to the requirement $\tau_1 = \tau_2 = \sigma_4 = \sigma_5 = 0$, and rather small failures to conform suffice to account for their observed posterior probabilities. To illustrate: In Table III of their paper, *S* AH, for one of the strongest signals (-1.85 log ml) gives $\Omega_5 = 1/9$ and $\Omega_4 = 1/4$. In this experiment, $\Omega = 1$, so if $\sigma_5 = 0$ per instructions, then

$$1/9 = \Omega_5 = \frac{1-p}{1-q}.$$

Suppose that $q = 0$ (HT). Then $p = 8/9$, and from

$$1/4 = \Omega_4 = \frac{(8/9)\sigma_4 + (1/9)\tau_4}{\tau_4},$$

we have $\sigma_4/\tau_4 = 5/32$. Thus, if τ_4 were .32, σ_4 need only be .05 to account for the data. If q were increased (or τ_4 decreased), σ_4 would be even smaller. The vast majority of the differences between Ω_4 and Ω_5 in Nachmias and Steinman's data are even easier to account for by small deviations from the assumption $\sigma_4 = \sigma_5 = 0$. In fact, the data in Nachmias and Steinman's Table III can be fit by HT (i.e., by assuming $q = 0$), using the general probabilistic response model of Matrix 1 and allowing the values of p, σ_i, τ_i to vary between rows of their table. (Since each row has data from separate experimental sessions, this variation of parameters is quite reasonable.) As will be shown subsequently, these data cannot be fit by 2-LT (with $q > 0$). This is not to say that HT is correct for absolute visual detection, but merely that their data do not suffice to disprove it.

As for Sperling's (1965) data, they are quite consistent with either HT or 2-LT. By Equation 2, $P(i|SN)$ is bound to increase with p (hence, with signal energy) for $\sigma_i > \tau_i$ and to decrease with signal energy for $\sigma_i < \tau_i$. If σ_i/τ_i increases with increasing confidence that the signal was presented, then the distribution of energies for which a given category is used is bound to shift toward higher energies for increasing confidence.

Divergence between 2-LT and SDT for CR Data

We have already noted one clear-cut way to reject 2-LT, pointed out by Broadbent (1966). According to 2-LT, the CR-ROC curve must either be two-limbed rectilinear, or must indicate poorer performance than "yes-no" detection. A curvilinear CR-ROC curve which lies at or above "yes-no" ROC points rules out 2-LT. However, existing data fail to reject 2-LT by this test.

A related point is that, if the probabilistic response matrix (Matrix 1) is manipulated by payoffs, instructions, or signal probabilities, then the resulting ROC points (Equation 3) need not lie on a single-valued curve. To illustrate this, consider the following numerical example, with $n=3$, $p = .8$, $q = .2$, and two different versions of Matrix 1, corresponding to different hypothetical instructions and/or payoffs:

	I				II		
D	.8	.2	0	D	.4	.4	.2
$\bar{D}$	0	.2	.8	$\bar{D}$	.1	.4	.5.

For I, by Equation 3, $x_1 = .16$, $y_1 = .64$, while for II, $x_1 = .16$, $y_1 = .34$. That is, the false-alarm rate .16 corresponds to hit rates of .34 or .64, depending on conditions.

This prediction might be difficult to test, since negative results might result from failure to induce experimentally such different response matrices as I and II. (The probabilities σ_i, τ_i are, of course, not directly observable.) Nevertheless, any serious noninvariance of ROC curves, either between "yes-no" and CR methods, or between different CR motivational conditions, would force a modification of SDT in the direction of probabilistic response models (of either a discrete-state or continuous type). Nachmias (1968) failed to find shifts in CR-ROCs.

The sharpest divergence of 2-LT from SDT, for CR data, lies in the prediction of an upper bound on posterior odds of the signal, given the response. By Equation 4, for any values of σ_i, τ_i, if $1 > p > q > 0$, then

$$\infty > p/q \geq \Omega_i/\Omega \geq (1-p)/(1-q) > 0. \qquad [5]$$

That is, the posterior probability of the signal cannot exceed an upper bound of $p/(p+q)$ or be smaller than $(1-p)/(2-p-q)$. On the other hand, SDT predicts no upper bound on Ω_i/Ω when the signal-plus-noise variance, σ_{SN}^2, is at least as large as the noise variance, σ_N^2.

Figure 2 shows values of Ω_1 that are quite high. Since $\Omega = 1$ for those data, $P(SN|1) = .95$ corresponds to $\Omega_1/\Omega = 19$ or $q < .05$. Very low values of q are uncomfortable for 2-LT, since they yield poor fits to "yes-no" ROC curves for low or moderate signal energy. On the other hand, the values of Ω_6 in Figure 2 are quite comfortable [$P(SN|6) = .21$ corresponds to $\Omega_6/\Omega = .27$, or $p > .73$, when q is small].

Similarly, the recognition memory data of Norman and Wickelgren (1965) also show high values of $P(SN|1) = P(\text{old item}|1)$, contradicting 2-LT (although, as pointed out above, their argument leading to rejection of 2-LT is erroneous). Their values of posterior probability for low-confidence, on the other hand, seem to be asymptotic at about .20.

The Nachmias and Steinman data are devastating for 2-LT, since 12 out of 18 estimates of $P(SN|1)$ are 1.00, corresponding to $\Omega_1 = \infty$ or $q = 0$. For S RS, seven out of eight of the estimates of Ω_1 are ∞, and the eighth is 13.3; the infinite values cut across a range of signal strengths. These results reject 2-LT out of hand. The case of 2-LT with $q \simeq 0$ reduces to HT, which is well known to be false for absolute visual detection (Swets, Tanner, & Birdsall, 1961).

Nachmias and Steinman say, in their discussion:

> the difficulties of two-state threshold theory are due in large measure to the auxiliary assumption that observers can be induced by instructions to use some response categories only when threshold is exceeded and others only when it is not. Conceivably, the theory would fare much better if this particular assumption is abandoned in favor of some other kinds of constraints . . . . However, without some constraint on . . . [the σ_i and τ_i . . .], the two-state threshold theory probably is untestable altogether by rating scale or "yes-no" experiments . . . [p. 1212–1213].

Thus, Nachmias and Steinman were well aware of the critical role played by the auxiliary assumption that $\tau_1 = \tau_2 = \sigma_4 = \sigma_5 = 0$ in their tests of 2-LT, but apparently missed the fact that their data reject 2-LT regardless of any constraints on the σ_i and τ_i.

In general, 2-LT may be rejected by estimating p and q from "yes-no" and/or forced-choice detection data (Luce, 1963a) and then showing that the estimated bounds of Inequality 5 are violated in CR data.

To summarize this section, the standard arguments against 2-LT, based on CRs, have been shown to be based on an unwarranted specialization of the response model (Matrix 1). Nevertheless, 2-LT does diverge markedly in two ways from SDT, for CR data. One divergence, involving bounds on Ω_i/Ω, leads to clear rejection of 2-LT for absolute visual detection. The second, involving noninvariance of the CR-ROC, has not been adequately tested.

Second Choices and 2-LT

According to HT, second choices in m-alternative forced-choice detection must be random. This is because at most one alternative (the one with the signal) can produce State D, and this alternative, if such exists, will be chosen first. The second choice is necessarily from among $\bar{D}$ alternatives, and can only be random. Various second-choice data contradict this prediction of HT (Green & Swets, 1966, pp. 108–110).

The fact that second-choice data are qualitatively consistent with low-threshold theories was recognized by Swets (1961), and has not been a source of much confusion. For completeness, the argument is reiterated here. For 2-LT, D can occur for more than one alternative, hence, the second choice can on some trials be made from among D alternatives and thus be nonrandom. Careful quantitative comparisons of 2-LT and SDT in this respect have not been made.

2-LT and Psychometric Functions

One finding injurious to 2-LT is that the estimated values of $q = P(D|N)$ decrease with increasing signal energy (Green & Swets, 1966, p. 144; Nachmias & Steinman, 1963; Swets, 1961). On the face of it, $P(D|N)$ should not depend on properties of the signal, and should remain constant as $P(D|SN)$ varies.[3]

[3] R. Duncan Luce (personal communication, July 1968) suggests an explanation of the relation between $P(D|N)$ and $P(D|SN)$ based on the neural quantum theory. The location of the boundary between $\bar{D}$ and D may be at 0, 1, 2, $\cdots$ quanta, and may vary as a function of sensory conditions, including signal strength. Thus, $P(D|N)$ might decrease as $P(D|SN)$ increases, if different signal energies are presented in separate blocks of trials.

A second argument involving variation of signal energy is the following. Consider an experiment where a complete "yes-no" psychometric function is obtained at fixed false-alarm rate, by interspersing signals of varying energy, including zero energy. Let $\mu = P(\text{"yes"}|D)$. The false-alarm rate is at least μq, if 2-LT is correct. Thus, very low false-alarm rates can only be achieved if μ is considerably less than one (assuming q is appreciably above zero; even for $q = .05$, a false-alarm rate of .01 can only be achieved if $\mu \leq .20$). But the hit rate cannot exceed $1 - (1 - \mu)p$, and thus the psychometric function must reach an asymptote less than or equal to μ, as p approaches one.

This prediction does not, of course, apply to experiments where a single block of trials contains signals of constant energy interspersed with blank trials, for there, μ can vary from block to block, hence, from energy to energy.

Barlow (1956) published two complete psychometric functions for absolute visual detection, with a single false-alarm rate for each function. The estimated false-alarm rates (300 blank trials) were .00 and .01, and the corresponding maximum (though not yet asymptotic) hit rates were about .94 and .97. (Barlow actually used a 3-category CR method, but the above argument applies perfectly well to the "yes" and "yes-or-possible" responses separately.) These data seem clearly inconsistent with 2-LT unless q is nearly zero; but $q \simeq 0$ again reduces to HT, which is false for absolute visual detection (Swets et al., 1961).

It seems intuitively quite likely that for very intense signals, the hit rate will approach 1.0, no matter how low the fixed (but nonzero) false-alarm rate is. This intuition suggests a modification of 2-LT: for intense signals, S enters a third state, denoted by D^*, in which he is certain that the signal was presented.

Note that the third state must be separated from $\bar{D}$ and D by a high threshold, that is, $P(D^*|N)$ must be approximately zero. For if $P(D^*|N) = q' > 0$, then again, the false-alarm rate exceeds $q' \cdot P(\text{"yes"}|D^*)$, and the same argument as above shows that psychometric functions cannot reach one for low false-alarm rates. The high-threshold state D^* also promises to account for the Nachmias and Steinman data, since $P(SN|1)$ can be 1.0 if $P(1|D^*) > 0$, $P(1|D) = P(1|\bar{D}) = 0$.

Three-State Low- and High-Threshold Theory

The three-state theory sketched in the preceding paragraph has two thresholds, a low one between $\bar{D}$ and D (since $P(D|N) > 0$), and a high one between D and D^* (since $P(D^*|N) = 0$). It will be abbreviated 3-LHT. It is actually a specialization of Green's two-threshold theory (Swets, 1961); in the latter $P(D^*|N)$ is not restricted to be zero. (Green's theory might be denoted 3-LLT.) 3-LHT differs sharply from Atkinson's three-state "variable sensitivity" theory (1963) in which $P(\bar{D}|SN)$, $P(D|SN)$, and $P(D^*|SN)$ are not fixed but vary from trial to trial depending on reinforcements. Other multistate theories, based on neural quantum theory, have been treated by Luce (1963a, 1963b) and Norman (1964).

The sensory model for 3-LHT is summarized in the following matrix:

	$\bar{D}$	D	D^*
N	$1-q$	q	0
SN	p_0	p_1	p_2

where $p_0 + p_1 + p_2 = 1$. There are thus three independent sensory parameters. Since SN should increase the likelihood of D relative to $\bar{D}$, p_1/p_0 is assumed $> q/(1-q)$.

The response model for "yes-no" detection, for 3-LHT, is a direct generalization of Luce's model for 2-LT. To derive this generalization from likelihood ratio considerations, note that the sensory states $\bar{D}$, D, D^* correspond to signal-noise likelihood ratios of $p_0/(1-q)$, p_1/q, and ∞, respectively. On a likelihood ratio scale, there are five positions for a criterion: less than $p_0/(1-q)$, at $p_0/(1-q)$, between $p_0/(1-q)$ and p_1/q, at p_1/q, and above p_1/q. Where the criterion falls at a likelihood ratio with positive probability of occurring, the decision rule can include any probability mixture of "yes" and "no" for that value of likelihood ratio. Hence, there are five types of criterion, defined, with corresponding hit and false-alarm rates, in Table 1.

The ROC curve thus begins at $(0,p_2)$. Varying $\beta = P(\text{"yes"}|D)$, it traces a straight line segment to $(q, p_1 + p_2)$. Varying $\alpha = P(\text{"yes"}|\bar{D})$, it traces a straight line segment from $(q, p_1 + p_2)$ to $(1,1)$. Another way to look at this ROC curve is in terms of the Neyman-Pearson criterion: The decision rules of Table 1 maximize hit rate for a given false-alarm rate, under the sensory conditions of 3-LHT. (A comparable derivation for 3-LLT yields a bottom limb from $(0,0)$ to $(P(D^*|N),p_2)$ on the ROC curve.) The present theory is more general than 2-LT, to which it reduces when $p_2 = 0$.

The extra freedom of 3-LHT over 2-LT, through introduction of an extra parameter p_2, can be gotten rid of, when signal strength is varied within the experiment, by demanding that q be invariant across signal strengths. It then seems logical to

TABLE 1

LIKELIHOOD RATIO CRITERIA FOR 3-LHT, WITH CORRESPONDING HIT AND FALSE-ALARM RATES

Criterion	Hit rate	False-alarm rate
$c > p_1/q$	p_2	0
$c = p_1/q$, $P(\text{"yes"}|c) = \beta$	$p_2 + \beta p_1$	βq
$p_1/q > c > p_0(1-q)$	$p_2 + p_1$	q
$c = p_0/(1-q)$, $P(\text{"yes"}|c) = \alpha$	$p_2 + p_1 + \alpha p_0$	$q + \alpha(1-q)$
$p_0/(1-q) > c$	1	1

Note.—For Criterion c, S says "yes" for any state with likelihood ratio of SN to N greater than c, and with specified probability when the likelihood ratio equals c.

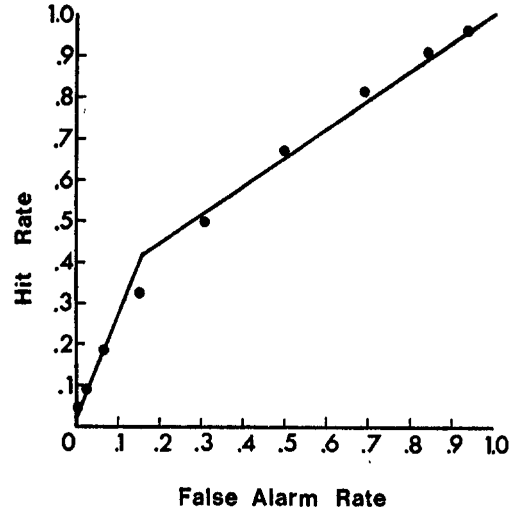

FIG. 3. The points are from an SDT ROC curve, with hypothetical N and SN distributions, whose parameters are given in Columns N and S_1 of Table 2. (The points represent equally spaced criteria on the underlying decision axis. The lines are a 3-LHT ROC curve, fit by eye to the points, with parameters given in Columns N and S_1 of Table 2.)

compare 3-LHT ROC curves, for different signal parameters p_1, p_2, with normal-unequal variance ROC curves generated by different values of $\mu_{SN} - \mu_N$ and $\sigma_{SN} - \sigma_N$. Such a comparison is presented in Figures 3–7. The nine points of each ROC curve are theoretical points from different normal-unequal variance ROC curves, generated by assuming $\sigma_{SN} - \sigma_N = (1/4)(\mu_{SN} - \mu_N)$ (see Green & Swets, 1966, pp. 96–98 for a discussion of this assumption, which is often used in fitting normal-unequal variance curves). The noise distribution and the five signal distributions for Figures 3–7 are all normal, with parameters given in the first two rows of Table 2. The 3-LHT ROC curves fitting these points were constrained to have $q = .16$; the values of p_0, p_1, p_2 are

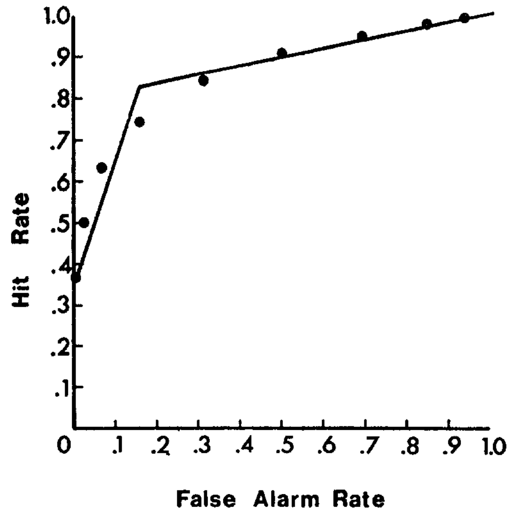

FIG. 5. Same as Figure 3, except that points and lines are based on Columns N and S_3 of Table 2.

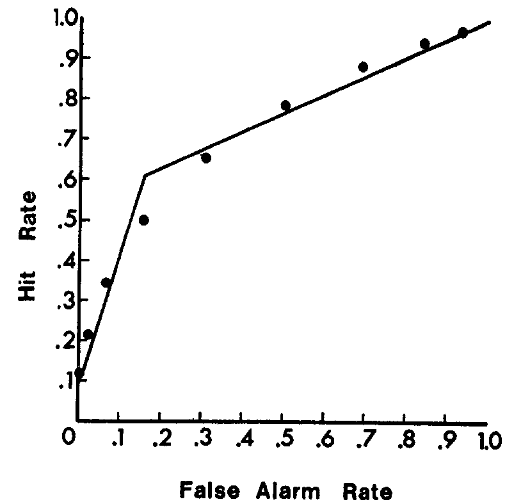

FIG. 4. Same as Figure 3, except that points and lines are based on Columns N and S_2 of Table 2.

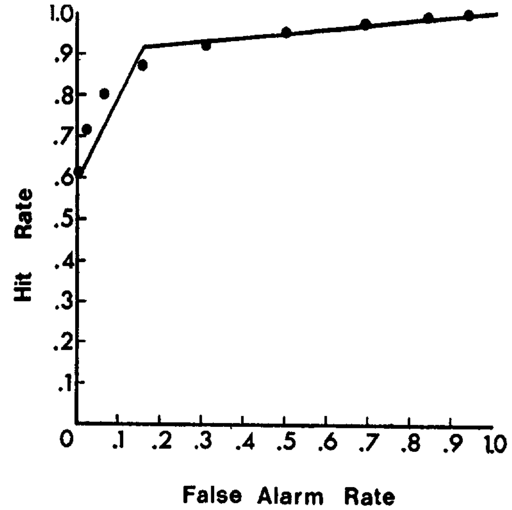

FIG. 6. Same as Figure 3, except that points and lines are based on Columns N and S_4 of Table 2.

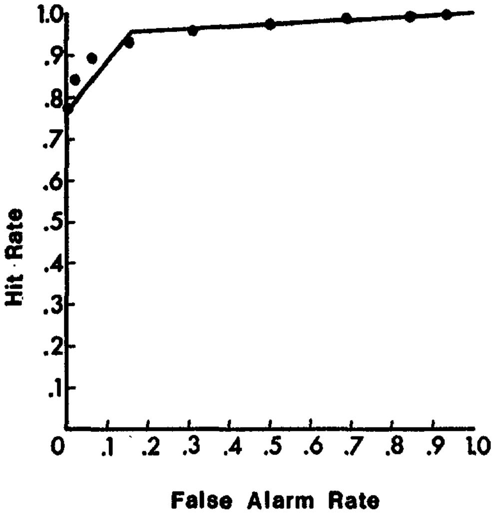

FIG. 7. Same as Figure 3, except that points and lines are based on Columns N and S_5 of Table 2.

presented in the last three rows of Table 2 (for N, the table gives $1 - q$, q, 0 instead). The rectilinear ROC curves were fit by eye. Because of binomial variability in estimates of hit and false-alarm rates, very careful measurements will be needed to distinguish the two-parameter family of 3-LHT ROC curves from the widely used normal-unequal variance curves. The fit of the 3-LHT curves, with two parameters free, is not necessarily better than that of 2-LT to the same points, but varying p_1, p_2 as a function of signal energy is much more reasonable than varying q, which would be required for comparable fits. Moreover, just as SDT cuts down on freedom of curve fitting by use of the empirically justified relation $\sigma_{SN} - \sigma_N = (1/4)(\mu_{SN} - \mu_N)$, so

TABLE 2

PARAMETERS FOR NORMAL NOISE AND SIGNAL DISTRIBUTIONS USED TO GENERATE THE POINTS OF FIGURES 3–7, AND CORRESPONDING VALUES OF p_0, p_1, p_2 FOR THEORETICAL CURVES OF 3-LHT

Item	N	S_1	S_2	S_3	S_4	S_5
μ	0.000	0.500	1.000	2.000	3.000	4.000
σ	1.000	1.125	1.250	1.500	1.750	2.000
p_0	$1 - q = .84$	.58	.39	.17	.08	.04
p_1	$q = .16$	.40	.51	.47	.32	.20
p_2	0	.02	.10	.36	.60	.76

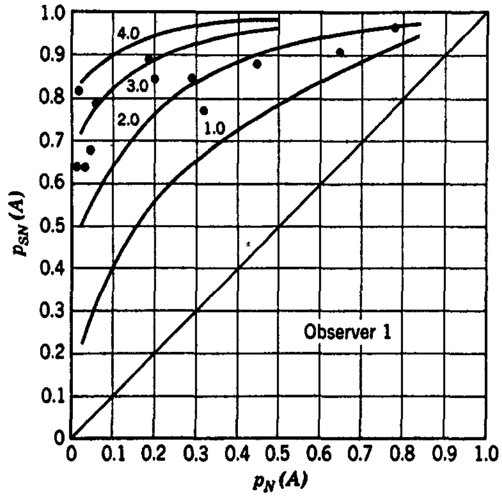

FIG. 8. Results of a "yes-no" visual detection experiment by Swets, Tanner, and Birdsall (*1961*), for Observer 1. (Points are estimated hit and false-alarm rates, each based on 200 trials in one experimental session. Payoffs varied between points. The theoretical curves represent different values of $(\mu_{SN} - \mu_N)/\sigma_N$, with $\sigma_{SN} - \sigma_N = (1/4)(\mu_{SN} - \mu_N)$. Reprinted with permission of the American Psychological Association.)

we may hope to find a similar relation among p_0, p_1, and p_2. (A first approximation to the values in Table 2 is given by $\sqrt{p_0} + \sqrt{p_2} = 1$, but this is not proposed for serious consideration. Perhaps theoretical interpretations of $\bar{D}$, D, D^* will one day yield a rational relation.)

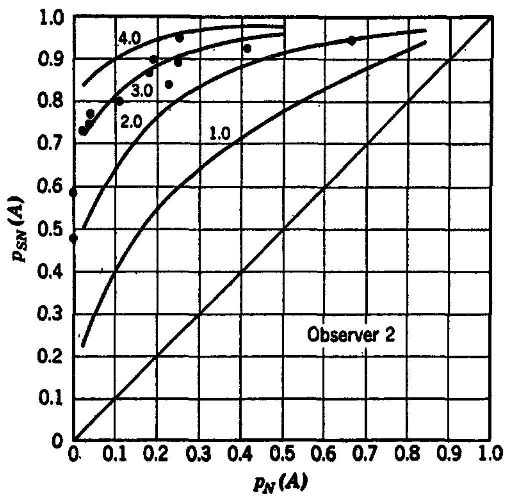

FIG. 9. Same as Figure 8, for Observer 2.

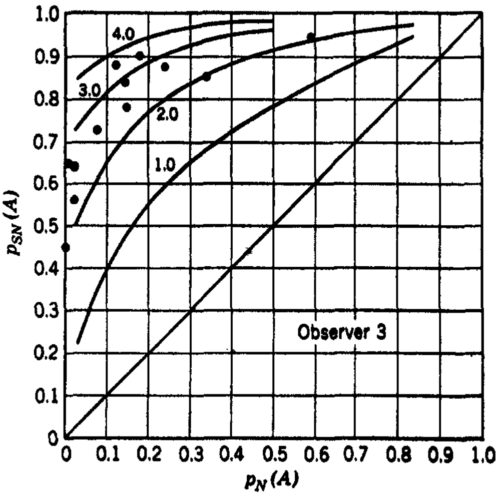

FIG. 10. Same as Figure 8, for Observer 3.

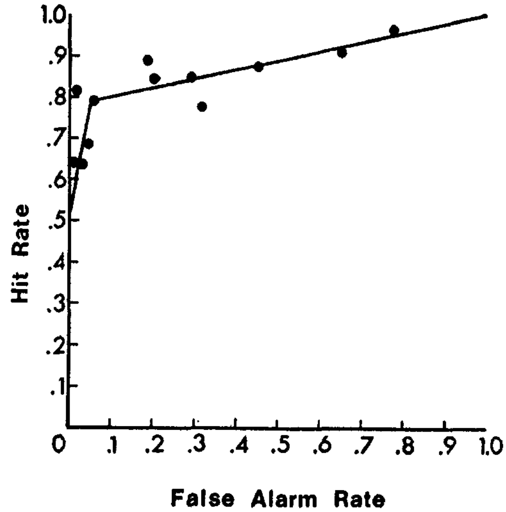

FIG. 12. The points of Figure 8 (Observer 1) are reproduced, with a 3-LHT ROC curve drawn by eye to fit them.

A further comparison of normal-unequal variance and 3-LHT may be obtained by examining typical empirical ROC curves. For this purpose, Figure 8 of Swets, Tanner, and Birdsall (1961) has been reproduced as Figures 8–11 of the present paper. These data are based on a visual absolute detection experiment, with a signal of .78 footlamberts. Each point is based on a session with 200 *SN* and 200 *N* trials, with different points representing different payoffs. In Figures 12–15, the same data points are reproduced, with theoretical curves of 3-LHT fitted by eye to them. In this case, no constraint was put on the three parameters of 3-LHT across the four curves, since the data are from four different observers.

The fit of 3-LHT to these data is about as good as could be expected, given the binomial variance of the points. To see this, note that most of the largest deviations are produced by points that would also deviate from *any* monotonic function, and thus are almost surely far from the "true" false-alarm–hit rate combinations for the corresponding conditions. There is no systematic deviation of the points from the

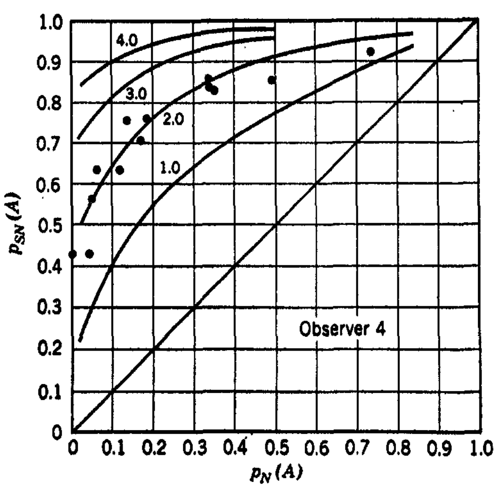

FIG. 11. Same as Figure 8, for Observer 4.

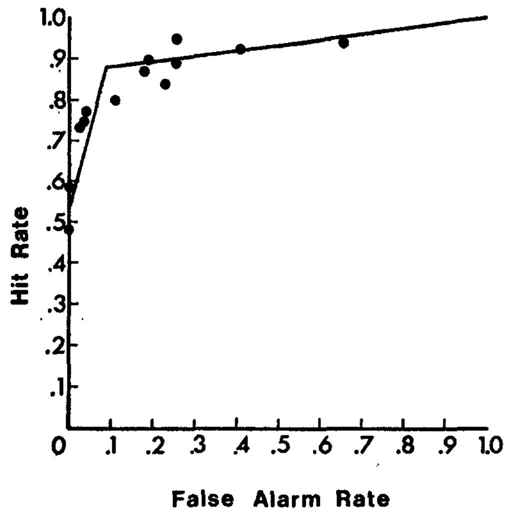

FIG. 13. Same as Figure 12, for Observer 2.

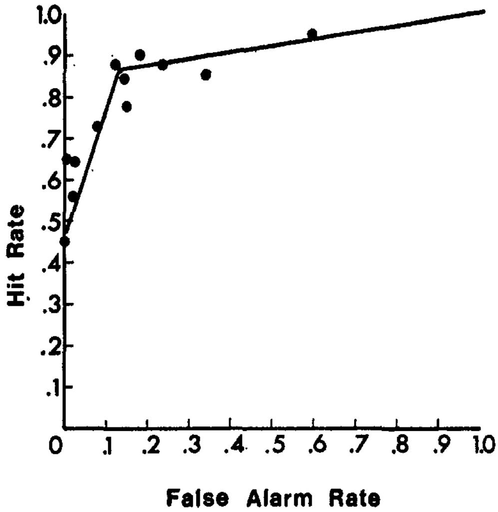

FIG. 14. Same as Figure 12, for Observer 3.

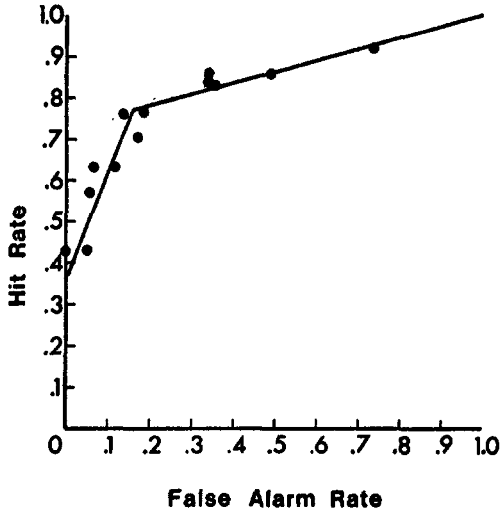

FIG. 15. Same as Figure 12, for Observer 4.

curves except perhaps for Observer 2. On the other hand, the points do seem to deviate systematically from both the normal-unequal variance curves and from 2-LT curves (see Luce, 1963a, Figure 2 for 2-LT fits to the same data). In fact, the points for high false-alarm rate tend to fall near curves with lower values of $\mu_{SN} - \mu_N$ than do points with low false-alarm rate. This is exactly the kind of systematic deviation that invites a better fit by the straight upper limb of 3-LHT. The points do not tend (except, perhaps, for Observer 2) to "float free" of the corner, as they do when 2-LT ROCs are fitted. Also, the presence of several points with estimated 0 false-alarm rate does nothing to harm the fit of 3-LHT relative to SDT and 2-LT, since only 3-LHT can account for true false-alarm rates of 0 with positive hit rates.

We have already seen that 3-LHT provides plausible ROC curves, without necessarily varying $P(D|N)$, as $P(D|SN)$ varies with changing energy. It should be pointed out in addition that this theory avoids other criticisms of 2-LT.

For one thing, the psychometric function can approach one, with low false-alarm rate, merely by letting p_2 approach one. This was the original motivation for the theory.

Second, 3-LHT has at least as much flexibility as 2-LT in handling CR data, but does not impose the upper bound $p/q \geq \Omega_i/\Omega$ of Inequality 5. In fact, the analog to Matrix 1, for 3-LHT, is:

$$\begin{array}{c|cccc} & 1 & 2 & \cdots & n \\ \hline D^* & \rho_1 & \rho_2 & \cdots & \rho_n \\ D & \sigma_1 & \sigma_2 & \cdots & \sigma_n \\ \bar{D} & \tau_1 & \tau_2 & \cdots & \tau_n \end{array} \qquad [6]$$

where

$$\sum_{i=1}^{n} \rho_i = \sum_{i=1}^{n} \sigma_i = \sum_{i=1}^{n} \tau_i = 1.$$

By the same reasoning as that leading to Equation 4, we have

$$\frac{\Omega_i}{\Omega} = \frac{p_2\rho_i + p_1\sigma_i + p_0\tau_i}{q\sigma_i + (1-q)\tau_i}. \qquad [7]$$

If σ_i, τ_i approach zero, then Ω_i/Ω becomes as large as we like. The lower bound remains, however; using the inequality $p_1/p_0 > q/(1-q)$, we get

$$\Omega_i/\Omega \geq p_0/(1-q). \qquad [8]$$

Thus, the large values of Ω_1 in Figure 2, and the even larger ones of Nachmias and Steinman, offer no difficulty for 3-LHT.

We conclude that the problems besetting 2-LT are avoided by 3-LHT. Note that they cannot be avoided without introducing a high threshold; if $P(D^*|N) = q' > 0$, then $p_2/q' \geq \Omega_i/\Omega$, etc.

Several reasonable restrictions on the response model Matrix 6 are worth mentioning. We may expect that ρ_i/σ_i and σ_i/τ_i will both be decreasing functions of i. Moreover, most of the ρ_i—perhaps, for small n, all but ρ_1—may be zero. That is, given the existence of the high-threshold state D^*, S may use exclusively high, or highest confidence when he is in that state. For very high energies of signal, and low number n of responses, S is apt to use Response 1 all the time, hence $\rho_1 = 1$, $\rho_2 = \cdots = \rho_n = 0$. Whether the response matrix, Matrix 6, is held constant with changing signal energy, or is allowed to vary, depends of course on whether different signal energies are randomly interspersed or are segregated into distinct blocks.

What does 3-LHT predict regarding the comparison of "yes-no" and CR-ROC curves? By procedures analogous to those used earlier, we can show that a necessary and sufficient condition for all the CR points to lie on the "yes-no" curve is that Matrix 6 takes the form

$$\begin{array}{c|cccccccc} & 1 & 2 & \cdots & k & k+1 & \cdots & n \\ \hline D^* & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ D & 0 & \sigma_2 & \cdots & \sigma_k & 0 & \cdots & 0 \\ \bar{D} & 0 & 0 & \cdots & 0 & \tau_{k+1} & \cdots & \tau_n. \end{array}$$

On the other hand, if the response matrix departs from this form, then the inequality $p_1/p_0 > q/(1-q)$ implies that CR performance is strictly worse than "yes-no." Thus, it is quite plausible to try to reject 3-LHT by showing curvilinear CR-ROC curves that match "yes-no" performance.

A Possible Method for Rejecting General Threshold Theories

The lower bound (Inequality 8) on posterior probability is not peculiar to 3-LHT but is a feature of any threshold theory worthy of the name, including the low-threshold theory of Swets, Tanner, and Birdsall (1961) with a continuum of states above threshold. To prove this, let A be the set of all states in the sensory model, let R be a response, and let

$$\Omega_R = P(SN|R)/P(N|R)$$

be the posterior odds of signal given Response R. Let $P(R|a)$ be the probability of Response R in State a (if a deterministic response model is assumed, this will be either 0 or 1, for any given a in A). Then by Bayes' theorem we have:

$$\Omega_R/\Omega = P(R|SN)/P(R|N) = \frac{\sum_{a\epsilon A} P(R|a)P(a|SN)}{\sum_{a\epsilon A} P(R|a)P(a|N)}. \quad [9]$$

(Where appropriate, the sums over $a\epsilon A$ are replaced by integrals.) Let A_R denote the set of States a such that $P(R|a)$ is nonzero. Then it follows from Equation 9 that:

$$\sup_{a\epsilon A_R} \frac{P(a|SN)}{P(a|N)} \geq \Omega_R/\Omega \geq \inf_{a\epsilon A} \frac{P(a|SN)}{P(a|N)}. \quad [10]$$

Let B denote the greatest lower bound (possibly 0) of $P(a|SN)/P(a|N)$, for all possible $a\epsilon A$. For any $\delta > 0$, let A_δ be the set of all a such that $P(a|SN)/P(a|N) < B + \delta$. Then A_δ is nonempty, and S can choose his response criteria for some R so that A_R is included in A_δ. It then follows from Inequality 10 that $B+\delta \geq \Omega_R/\Omega \geq B$. We have thus proved that the greatest lower bound of Ω_R/Ω, over all possible response criteria, coincides with B, the greatest lower bound of $P(a|SN)/P(a|N)$.

For any threshold theory, the set A must include a below-threshold state, $\bar{D}$. For any above-threshold State $a \neq \bar{D}$, it is assumed that the likelihood of a, relative to that of $\bar{D}$, is greater given SN than given N. This is equivalent to:

$$\inf_{a\epsilon A} P(a|SN)/P(a|N) = P(\bar{D}|SN)/P(\bar{D}|N). \quad [11]$$

Furthermore, for any signal of only moderate strength, $P(\bar{D}|SN)$ must take some positive value p_0 (it would be a very peculiar threshold if all nonzero signals were always above threshold). Continuing to denote $P(\bar{D}|N)$ by $1 - q$, we have from Inequality 10 and Equation 11, for any R:

$$\Omega_R/\Omega \geq p_0/(1-q) > p_0 > 0. \qquad [12]$$

This is the same result as Inequality 8, generalized to arbitrary threshold theories.

One way to reject all threshold theories is to show that even for signals of moderate strength, Responses R can be found (say, in a confidence-rating experiment with suitable instructions) such that Ω_R/Ω takes on very small values. If we can show, in other words, that p_0 is very near zero for almost all signals, then the subthreshold State $\bar{D}$, if any, plays no significant role in detection.

As mentioned, most plots of Ω_R/Ω seem to show an asymptote clearly above zero. The only exception known to the present author is the study of Carterette and Cole (1963), although that study involves CR that a response was correct, rather than that a signal was present. At any rate, these data show the statistical feasibility of generating Ω_R/Ω very close to zero.

Unfortunately, failure to reject threshold theory by the test outlined above does not lead to acceptance of threshold theory and rejection of SDT. In SDT, with $\sigma_{SN} > \sigma_N$, $P(a|SN)/P(a|N)$ is again bounded below, and the bound is apt to be similar in magnitude to $p_0/(1 - q)$ as estimated from specific threshold theories. This shall be illustrated by comparing the lower bounds for 3-LHT, derived from the best fitting curves of Figures 3–7, with the lower bounds implied by SDT, using the normal-unequal variance assumptions that generated the points of Figures 3–7.

Assuming normal distributions, with $\sigma_N = 1$, $\sigma_{SN} = 1 + b$, $\mu_N = 0$, $\mu_{SN} = 4b$,

$$P(a|SN)/P(a|N) = \frac{1}{1+b} e^{-(1/2)\left\{\left(\frac{a-4b}{1+b}\right)^2 - a^2\right\}}.$$

Differentiating the above with respect to a, setting the derivative equal to zero, we can solve for a at which a minimum is achieved. Substituting this value of a back in Inequality 12, we obtain the SDT lower bound on $P(a|SN)/P(a|N)$:

$$\frac{1}{1+b} e^{-\frac{8b}{2+b}}.$$

Values of this lower bound, for different values of b, are compared in Table 3 with corresponding values of $p_0/(1 - q)$ for 3-LHT. It is evident from this comparison that any data failing to reject 3-LHT on the basis of lower bounds, cannot convincingly favor 3-LHT over SDT.

This does not mean, of course, that the attempt to reject *all* threshold theories, by producing very low values of Ω_R/Ω, is doomed to failure. It seems quite worth trying. But such an attempt cannot by itself lead to acceptance of threshold theories rather than SDT.

TABLE 3

MINIMUM VALUES FOR Ω_R/Ω FOR NORMAL-UNEQUAL VARIANCE AND 3-LHT ROC CURVES OF FIGURES 3–7

Figure	$b = \sigma_{SN} - \sigma_N$	inf Ω_R/Ω SDT	inf Ω_R/Ω 3-LHT
3	.125	.557	.691
4	.250	.330	.464
5	.500	.136	.202
6	.750	.065	.095
7	1.000	.035	.048

REFERENCES

ATKINSON, R. C. A variable sensitivity theory of signal detection. *Psychological Review*, 1963, **70**, 91–106.

BARLOW, H. B. Retinal noise and absolute threshold. *Journal of the Optical Society of America*, 1956, **46**, 634–639.

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

CARTERETTE, E. C., & COLE, M. Repetition and confirmation of messages received by ear and by eye. *Quarterly Journal of Experimental Psychology*, 1963, **15**, 155–172.

EGAN, J. P., SCHULMAN, A. I., & GREENBERG, G. Z. Operating characteristics determined by binary decisions and by ratings. *Journal of the Acoustical Society of America*, 1959, **31**, 768–773.

GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.

LARKIN, W. D. Rating scales in detection experiments. *Journal of the Acoustical Society of America*, 1965, **37**, 748–749.

LEE, W. Choosing among confusably distributed stimuli with specified likelihood ratios. *Perceptual and Motor Skills*, 1963, **16**, 445–467.

LUCE, R. D. *Individual choice behavior*. New York: Wiley, 1959.

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

LUCE, R. D. Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. Vol. 1. New York: Wiley, 1963. (b)

MARKOWITZ, J., & SWETS, J. A. Factors affecting the slope of empirical ROC curves: Comparison of binary and rating responses. *Perception and Psychophysics*, 1967, **2**, 91–100.

NACHMIAS, J. Effects of presentation probability and number of response alternatives on simple visual detection. *Perception and Psychophysics*, 1968, **3**, 151–155.

NACHMIAS, J., & Steinman, R. M. Study of absolute visual detection by the rating-scale method. *Journal of the Optical Society of America*, 1963, **53**, 1206–1213.

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

NORMAN, D. A., & WICKELGREN, W. Short-term recognition memory for single digits and pairs of digits. *Journal of Experimental Psychology*, 1965, **70**, 479–489.

SHIPLEY, E. F. Detection and recognition with uncertainty: An experiment and a revised model. Unpublished doctoral dissertation, University of Pennsylvania, 1961.

SPERLING, G. Temporal and spatial visual masking. I. Masking by impulse flashes. *Journal of the Optical Society of America*, 1965, **55**, 541–559.

STEVENS, S. S., MORGAN, C. T., & VOLKMANN, J. Theory of the neural quantum in the discrimination of loudness and pitch. *American Journal of Psychology*, 1941, **54**, 315–335.

SWETS, J. A. Is there a sensory threshold? *Science*, 1961, **134**, 168–177.

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

WATSON, C. S., & Bourbon, W. T. Rating scales and two-state threshold models. *Journal of the Acoustical Society of America*, 1965, **38**, 667–668.

WATSON, C. S., RILLING, M. E., & BOURBON, W. T. Receiver-operating characteristics determined by a mechanical analog to the rating scale. *Journal of the Acoustical Society of America*, 1964, **36**, 283–288.

WICKELGREN, W. A. A test of Luce's two-state decision rule. *Psychonomic Science*, 1967, **9**, 91–92.

WICKELGREN, W. A. Testing two-state theories with operating characteristics and a posteriori probabilities. *Psychological Bulletin*, 1968, **69**, 126–131.

(Received June 3, 1968)