A NEURAL TIMING THEORY FOR RESPONSE TIMES AND THE PSYCHOPHYSICS OF INTENSITY [1]

R. DUNCAN LUCE   AND   DAVID M. GREEN[2]

Institute for Advanced Study, Princeton, New Jersey   *University of California, San Diego*

This psychophysical theory involves the following fundamental assumptions. At a hypothetical neural decision center, signal intensity is represented by several independent, parallel Poisson processes, whose rates are the same power function of physical intensity. All decisions about signal intensity are based on the observed times between successive neural pulses. The total number of these times observed per channel is at the option of the subject, up to the size of a memory store which is emptied when a decision is made. Overall response time is the sum of the decision latency, which depends both on the signal intensity and the decision rule, and a residual latency which is only assumed to be bounded. Decision rules are suggested for discrimination, recognition, magnitude estimation, detection, and simple reaction time designs, and predictions are derived from the theory in these cases and compared with existing data. Various familiar generalizations, such as Weber's and Bloch's laws and the inverse relation between reaction time and intensity, derive naturally from the theory. Crude estimates of all parameters—the exponent of the power function, the number of parallel channels, the size of the buffer store, and the bound on the residual times—are provided for sound intensity; estimates from different experimental designs appear to be reasonably consistent.

Elsewhere we have begun to evolve and test a detection and response time theory which, at least for pure tones of different intensities, seems to have merit (Green & Luce, 1967, 1971, 1972; Luce, 1966; Luce & Green, 1970). In this paper we formulate the theory more generally and show how it applies to a variety of psychophysical and reaction time (RT) designs.

The theory is based on three fundamental postulates which we now state; they are discussed more fully below.

1. The output of a sensory system, such as the ear, is treated as a set of sequences of discrete, brief events—"neural" *pulses*—on parallel channels with the following properties: (*a*) The duration of each pulse is negligible; (*b*) the times between successive pulses—the *interarrival times* (IATs)—on each channel and between channels are in-

dependent random variables (RVs); (*c*) when the intensity of a signal is constant, the IATs have a common distribution with an expected value that is a strictly monotonic decreasing function of the signal intensity.

We make two subsidiary assumptions which are conceptually quite secondary but are mathematically convenient and, apparently, are approximately correct. First, when the signal intensity is constant, these stochastic processes are assumed all to be *Poisson processes* with the same intensity parameter μ.[3] This means that on any one channel, the common distribution of the independent IATs is exponential and so, by 1*c*, its intensity parameter μ is a strictly increasing function of the signal intensity. Second, we shall assume that the latter function is, in fact, a power function. As we shall see, at least three different empirical results support this assumption.

2. Decisions as to when and/or what to respond are determined by decision rules that can be stated in terms of certain IATs.

[2] Requests for reprints should be sent to David M. Green, Department of Psychology, University of California, San Diego, La Jolla, California 92037.

[3] Undoubtedly, the values of the parameters differ in different channels. Our idealization is obtained by replacing the actual parameter values by average ones.

The principle underlying these rules (derived from 1c) is that any IAT is an estimate of (a monotonic function of) the signal intensity. We deal in this paper with only the simplest rules in which either a fixed number of IATs are observed or observations are continued until the first IAT occurs that is less than a criterion. Such models we call *timing models*. The time until a decision is reached is called the *decision latency*.[4]

One feature of the decision center is that it can store only a limited number of IATs at any one time. Sometimes, the observed behavior is limited by this storage capacity; other times it is limited by signal duration, which in effect limits the (expected) number of IATs.

3. The *residual latency*, which is defined as the overall RT less the decision latency, is an RV that is bounded and is independent of the decision latency.

Relation to Counting Models

The analysis of human sensory systems in terms of Poisson processes has become increasingly popular (Barlow, 1956, 1957; Bouman, 1961 [see references given there]; Creelman, 1961; Grice, 1968; Kohfeld, 1969; McGill, 1963, 1967; McGill & Goldberg, 1968; Siebert, 1965, 1968, 1970). The appeal is twofold: First, the observed all-or-none character of neural pulses invites description in terms of a continuous stochastic process. Second, of the possible processes, the Poisson is mathematically the most tractable and, as we argue below, is a tolerable approximation to some peripheral neural data. All of these other theories share the assumption that decisions are based not on IATs, but on a count of the Poisson events occurring within a prescribed time. The difference in our approach is exemplified by the problem of estimating an instantaneous heart rate. Either one can count the number of beats during a period of time and then calculate the average rate, or one can use the recip-

rocal of the time between the two successive beats spanning the instant of interest. The former estimate, which is usually based on a larger sample of IATs, is less variable, but because the average is constructed over some time period, it is not terribly sensitive to brief fluctuations in the heart rate. Since the detection of faint and/or brief signals that may arrive at any time is similar to noticing brief changes in the heart rate, we find the IAT decision rules appealing. In any event, timing models are worth investigating. Moreover, by emphasizing IATs rather than neural counters or accumulators, we avoid many of the elaborate and often ad hoc assumptions which are made concerning the memory for or integration of sensory information. It also has the distinct advantage of allowing one to derive the distribution of response times from the same assumptions as are used to arrive at the various stimulus-response probabilities, thereby permitting a more unified treatment of psychophysical and response time processes. Finally, certain empirical generalizations, such as Bloch's and Weber's laws, derive naturally.

Discussion of the General Theory

We repeat and discuss more fully our postulates. First, the sense organ, in our experimental studies the ear, is assumed to be a transducer that converts the intensity, $I(s)$, of the incoming fixed-frequency signal, s, into several trains of pulses with independent IATs. The intensity is represented by the temporal pattern of the pulses, not by their amplitude or spatial location; as a result, there is an inherent temporal ambiguity in knowing when changes of the input signal occur. In the case of constant intensity signals, we assume that the output on each of several channels is a Poisson process with intensity parameter $\mu(s)$ that, by assumption and confirmed by experiment, increases with $I(s)$. (This can be generalized to nonconstant intensities by assuming that the hazard function of the stochastic process is a strictly increasing function of $I(s)$.)

Since the pulses we are talking about are those on which decisions are based and

---

[4] We follow the convention of using the word "time" for any time that we can observe directly and "latency" for any time that is an unobservable construct of a theory.

hence are, presumably, located somewhere central in the brain, and since the properties most pertinent to our assumptions have only been carefully measured in neurons in the periphery, it is not clear whether the existing physiological data are relevant. According to Grossman and Viernstein (1961), Rodieck, Kiang, and Gerstein (1962), and Kiang (1965, especially pp. 95–101), successive IATs in the auditory nerves of cats are usually independent of each other, but their distributions are at best not quite exponential, and sometimes are decidedly not exponential. In the best cases, the observed distribution of IATs do not rise vertically at 0, but reach their peak value only after a few milliseconds, following which they are often well approximated by an exponential decay. The failure to find extremely short IATs is interpreted as re-fractoriness in the neurons. However, Rose, Brugge, Anderson, and Hind (1967; also see Brugge, Anderson, Hind, & Rose, 1969; Hind, Anderson, Brugge, & Rose, 1967; Rose, Brugge, Anderson, & Hind, 1969), using very refined techniques on the squirrel monkey, have shown that the distribution appears to be very nearly geometrically spaced at the period of the input signal. Thus, with a 1000-hertz (Hz.) tone, the IATs are approximately multiples of 1 millisecond (msec.) and have a distribution of the form $p^n q$. Similar frequency locked IAT data, with pulsed clicks and cats as subjects, were reported by Kiang (1965, Ch. 4). Of course, the exponential is the continuous approximation to the geometric and, provided that we do not care to discriminate as finely as 1 msec., it makes little difference which we use. Thus, although the Poisson assumption may not be quite correct, it seems to be a satisfactory first approximation in many cases. Ultimately, the exponential distribution can be replaced by, for example, a gamma distribution, but many fewer closed-form expressions will result. Two useful surveys of statistical analyses of single-unit activity are Kiang (1968) and Perkel and Bullock (1968).

As we said, these data may not be relevant to our postulate. It should be noted that if a number of independent continuous stochastic processes with independent IATs are added, as may happen in parts of the nervous system, the resulting process tends toward the Poisson independent of the distribution of the original IATs.

There are relatively few data bearing directly on the dependence of $\mu(s)$ on $I(s)$, in particular, on the constancy of μ with constant I and on the discontinuous change in μ when we change the intensity discontinuously from one constant value to another, as in most psychophysical experiments. Kiang's (1968) data suggest that μ decreases over a period of minutes with a constant intensity signal, but over seconds the effect is negligible. He has also observed fibers in which $\mu(s)$ increases for a while with $I(s)$ and then diminishes as $I(s)$ is further increased.

As we shall see, psychophysical data indirectly suggest that for constant intensity signals,

$$\mu(s) = \begin{cases} \alpha[I(s) - I_0]^\gamma, & \text{if } I(s) \geq I_0 \\ 0, & \text{if } I(s) < I_0, \end{cases}$$

where α and γ are constants. Psychophysical data permit us to estimate γ, but not α. To anticipate, the exponent γ is obtained by magnitude estimation methods. S. S. Stevens (1970) marshals evidence to support the view that neural firing rates and other physiological measures of neural activity are power functions of intensity.

The second major component of the theory is a decision center that processes the sequences of pulses in some fashion to decide when to respond and/or which response to make. Since the IATs are independent and, with a constant intensity signal, identically distributed (actually, exponentially distributed under the Poisson assumption), it is plausible to assume that all decisions can be formulated in terms of the values of certain IATs. Exactly how decisions are reached depends both on inherent features of the nervous system and on the exact experimental situation, including our instructions and information feedback. By imposing different extreme experimental conditions, various limitations of the nervous system will become

dominant, and we thus hope to discover what they are. Our tactic will be to generate a variety of models[5] as we consider the observable consequences of various decision rules. The common feature of all of the rules described in this paper, however, is that decisions are based on comparisons of means of IATs with each other or with a criterion; thus, the mean IAT plays a role somewhat comparable to the likelihood ratio RV in the theory of signal detectability (TSD; see Green & Swets, 1966).

The basic principle underlying all of the decision rules is that each IAT is an estimate of the signal intensity; this stems from the assumption that the expected value of the IAT is a monotonically decreasing function of the signal intensity (under the Poisson assumption, an IAT is an estimate of $1/\mu(s)$). With this principle in mind, three simple decision rules arise depending on the experimental design.

First, there are experiments in which the onset of a (possible) signal is well defined, and the problem is to identify which signal is being presented (included are yes–no detection, recognition, and magnitude estimation). In such designs, one or more IATs can be observed under the prior knowledge that they all arise from the same (Poisson) process. The rules that we shall examine have, at most, three free parameters: the number, J, of parallel channels (nerve fibers) over which intensity information is transmitted; the number, κ, of IATs per channel that a subject collects before initiating a response, and (sometimes) a criterion, β, against which the mean of the IATs is compared. When the subject is under considerable time pressure, for example, in certain RT experiments and, as we argue below, whenever the signals are both faint and brief, we shall assume $\kappa = 1$. When there is no time pressure, the value

of κ must be estimated. It is clear that were the subject free to make κ arbitrarily large, he could, at the expense of taking enough time, make arbitrarily fine discriminations, which empirically is impossible. This leads us to postulate the existence of a buffer store in which only a limited number of IATs can be held prior to a decision. The size of the buffer, K, can be estimated from psychophysical experiments using brief but intense signals, such as recognition and magnitude estimation experiments.

Complications arise because it is not always clear which decision rule to assume; even when we restrict ourselves to rules only involving IATs and we specify J and κ, several alternatives seem equally plausible. A particularly thorny problem is how to treat simple detection experiments when the signal is presented only for a brief time. Just how should pulses that occur outside the observation interval be handled? It is evident that the observer ignores those events that occur far outside the listening interval, but it is unclear how sharply the observer's timing information can be used to demark the listening interval. Furthermore, the effect of this uncertainty must vary with signal intensity.

In those experiments in which the signal is presented for some fixed duration, there is a real possibility that the subject simply samples some fixed interval in time and counts the pulses in that interval. This is the usual analysis given by the counting models. If this is the case, then latency information should reflect variation in the end of this interval and in other fixed delays, but should be independent of which signal was present in the interval. We work out two timing models. The first simply waits for a neural event to terminate the trial, even if that event occurs well outside the listening interval. The second model assumes that IAT information initiates a response if it occurs before a deadline, which is set at the end of the listening interval, and initiates a response at that time if not before. It may well be that details of the instructions and the payoffs of the experimental task determine

whether the counting or timing model is correct in these experiments with fixed signal durations. In many fixed signal duration experiments, with no time pressure, the counting models probably provide the more plausible theory.

Second, there are experiments in which two or more signals are presented at well-defined times during each trial (included are forced-choice and pair-comparison designs). A mean IAT can be observed for each signal interval, with some number $k = J\kappa$ of observations entering into each mean, and then the means compared in order, for example, to decide which interval contained the most intense signal.

Third, there are experiments in which neither the occurrence nor identity of the signal is at issue, but only its time of occurrence (included are RT designs with random foreperiods and the method of free response). Here the task of the subject is to notice, by observing IATs, when the signal intensity is changed,[6] and the natural rule is for the subject to establish a criterion β with which he compares a running average of κ successive IATs on J channels. When that running average becomes less than a criterion β, a detection response is initiated. By increasing κ, the reliability of the detection can be improved, but at the expense of delaying it long after the actual onset of the signal. As we shall see, such rules are difficult to study, and we work out (incompletely) only the case of $J = 1$, $\kappa = 1$ for the RT experiment with random foreperiods. A special argument is used to generalize the results to $J > 1$. Considerably more research is needed on this class of rules.

Of course, rules other than these three types are possible. One of the most obvious for some of the above designs are truncated sequential rules. We have not undertaken their study here primarily because of the great difficulty one has in obtaining explicit expressions for response time densities.

[6] By the dictionary definition, "detection" is the ideal term for this class of experiments, but that word has long been appropriated for such special recognition designs as yes–no and two-alternative forced-choice in which one signal is of zero intensity.

Nevertheless, they should be explored in the context of this theory. Two papers relevant to optimal detection for somewhat different stochastic processes are Bather (1967) and Shiryaev (1963).

For experiments in which the signal onset is clearly signaled in another modality, another type of rule is possible which takes advantage of a property peculiar to the Poisson process. This is the fact that the time from an arbitrary random event, such as the signal onset, to the next pulse of the process is an RV with exactly the same distribution as that of the IATs. Thus, decisions become more efficient if these times are combined with the IATs. The whole paper can be reconstructed using this type of rule. We elected not to use it for two reasons. First, if the process is not strictly Poisson, then the time to the first pulse after an event does not have the same distribution as that of the IATs and, in general, it would be a mistake to treat it in the same way. Second, life has evolved in a world in which the potential onset of a signal is rarely clearly marked in any way—most signals go unnoticed unless they attract attention to themselves—and so we would be surprised if a mechanism would have evolved just to take advantage of the extra information available in experiments having trials.

It should be noted that in all cases where we use a criterion β, we assume that it is maintained at a constant value throughout an experimental run. This is almost certainly wrong as the existence of trial-by-trial sequential effects argues. Our strategy is to assume that the variability in β is sufficiently small so that we can bury it in the variability of the IATs without too gross an error. Later, if certain decision rules seem viable, we will have to use sequential data to separate IAT and β variability.

The present decision rules are wholly inadequate to handle frequency discrimination. One possible approach is to suppose that different frequencies activate different, overlapping sets of parallel channels, and the subject attempts to identify from the observed IATs which subset is activated.

For low frequencies, another approach is suggested by the data of Rose et al. (1967), namely, that frequency can be inferred from the discrete pattern of the IATs, which are multiples of the period of the input signal, but subject to some random variation.

Since the pulses are strung out in time, no matter which rule we postulate some time is consumed in arriving at a decision. In those situations with a well-defined signal onset, we can observe a response time, part of which is contributed by the latency of the sensory decision process. Because the expected value of the IAT increases with decreasing signal intensity, the decision latency and so the observed response time is magnified by using weak signals. In a sense, weak signals serve as our "microscope" to explore these elementary, but basic decision mechanisms of the brain.

We use the letter "l" to denote the generic density of sensory decision latencies, sometimes adding subscripts to identify which signal was presented and which response was made. Most of the mathematical work in the theory involves using the assumed decision rule and the other postulates of the theory to derive expressions for the various decision densities. These formulas include as parameters such things as the Poisson intensity parameters of the sensory output and, where appropriate, J, κ, and β.

The third major component of the theory is the sum of all times taken to complete the response excluding the time absorbed by the decision process. This includes, at least, the times (a) needed to transduce the energy into pulses, (b) required for a pulse to travel from the transducer to wherever decisions are made, (c) taken in travel from there to the locus of the response, and (d) needed for the muscles to act. This list may not be exhaustive. The sum of all these nondecision times we call the residual latency.[7] We assume that these residual

latencies are continuous RVs, that they are independent of each other and of the decision latencies, and that they are bounded. We denote their (unknown) density by r, with a subscript referring to the response made if there are two or more. Clearly, changing either the mode of stimulation or the mode of response will change r. The boundedness assumption means that there is some $\tau > 0$ such that $r(t) = 0$ for all $t \geq \tau$.

We know and assume very little about r. We do not attempt to construct any theory of the residual times. From our point of view, they are a nuisance factor making it difficult to get at the decision process. Although postulating a bound on the residual latencies seems most plausible, it should be realized that it is sometimes crucial in our analysis. This bound, together with the exponential IATs, which are therefore unbounded, sometimes permits us to get a fairly clear look into aspects of the sensory decision process uncontaminated by the residual times. Were the residual times unbounded, we could not use the tails of observed distributions to estimate parameters.

If we denote by f_j the density of observed response times when response j is made to the signal, then by our independence assumption, it is the convolution of the decision and residual densities, l_j and r_j (see Figure 1), that is,

$$f_j(t) = \int_0^t l_j(x) r_j(t - x) dx. \qquad [1]$$

Because all of our densities are 0 for $t < 0$, the Fourier transform[8] of f_j is

$$F_j(\omega) = \int_0^\infty e^{-i\omega t} f_j(t) dt,$$

$$\omega \text{ is real and } i^2 = -1.$$

---

a decision to a response. In talking with others, we discovered that this was badly misleading, especially for designs with trials, and so it seems best to alter the terminology. The residual latency corresponds to what was called the "irreducible reaction time" in the classical RT literature.

[8] We follow the convention of using small Latin letters for densities and their capitals for the corresponding Fourier transform.

[7] In our previous publications we called it the "response latency" because, in thinking about free response designs, we had focused on the time from
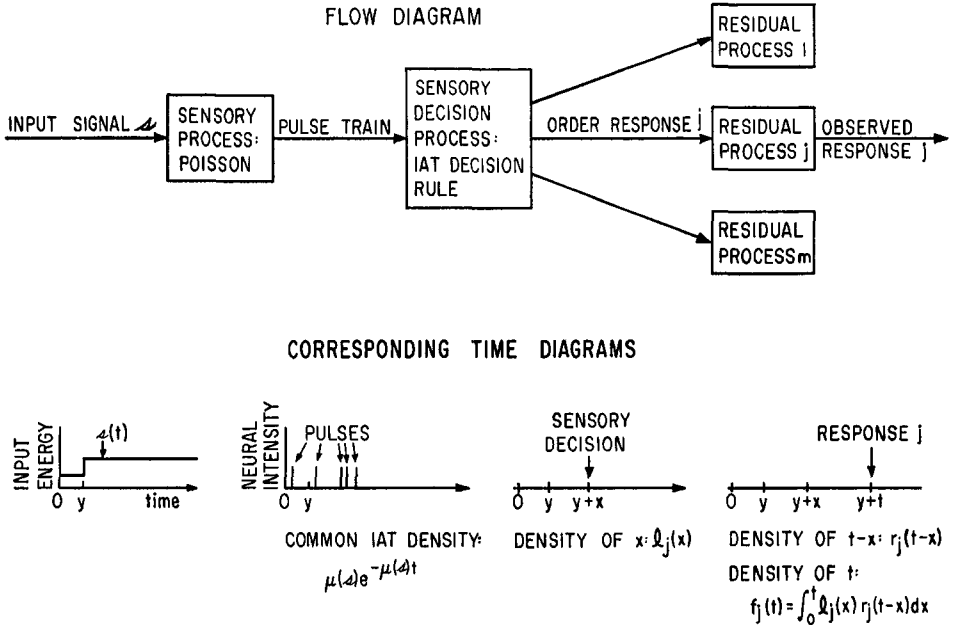
FLOW DIAGRAM

CORRESPONDING TIME DIAGRAMS

FIG. 1.   Block diagram of model and time diagrams of the output at each stage.

It is well known that if f_j is a convolution of l_j and r_j, then[9]

$$F_j(\omega) = L_j(\omega)R_j(\omega). \qquad [2]$$

This suggests the following general procedure for testing a model. State the proposed sensory decision rule and, suppressing the response label j, derive L (usually, though not always, by deriving l and then computing its Fourier transform) as an explicit expression in terms of parameters, such as the Poisson parameters of the sensory process, J, κ, and β. Devise ways to estimate these parameters (we often use the boundedness of r to show that the shape of the tail of f depends only on l and not on r, which permits us to use the tails to obtain estimates). Calculate from the data $F(\omega)/L(\omega)$, which is an estimate of $R(\omega)$. By inverting this transform, we have an estimate of r. We can then see, first, whether it is a density and, if so, whether it is independent of manipulations

of the parameters. If not, then doubt is cast on the decision rule. This was done successfully in Green and Luce (1971).

We show precisely how the boundedness of r can be used in parameter estimation. Suppose it has been shown that for sufficiently large t, l acts like a gamma distribution, that is,

$$l(t) \cong ct^m e^{-\eta t} \quad \text{for } t \geq \tau^*. \qquad [3]$$

Then by Equation 1,

$$f(t) = \int_{t-\tau}^{t} l(x)r(t-x)dx$$

$$[\text{since } r(t-x) = 0 \text{ for } t-x \geq \tau]$$

$$\cong c \int_{t-\tau}^{t} x^m e^{-\eta x} r(t-x)dx,$$
$$\text{for } t \geq \tau^* + \tau$$

$$= c \int_{0}^{\tau} (t-y)^m e^{-\eta(t-y)} r(y)dy$$

$$= e^{-\eta t} \sum_{i=0}^{m} a_i t^i, \qquad [4]$$

where

$$a_i = c\binom{m}{i}(-1)^{m-i}\int_{0}^{\tau} y^{m-i} e^{\eta y} r(y)dy.$$

[9] Other classical transforms, including the Laplace, have the same property of converting convolutions into multiplications. We use the Fourier because there exists an excellent computer program, known as the Fast Fourier Transform, for numerically calculating transforms and their inverses.

For $m = 0$, f is proportional to $e^{-\eta t}$ if $t \geq \tau^* + \tau$, and so η can be estimated from the tail of f. For $m > 0$, the dominant term is $t^m e^{-\eta t}$ and again η can be estimated. For some models, but not all, l is shown to have the above form.

Another approach was suggested many years ago by Christie and Luce (1956). Suppose the experiment is run twice with signals of different intensities, but with the same response. Assuming that the residual latency is unaffected by the signal intensity, we have

$$\frac{F_1(\omega)}{F_2(\omega)} = \frac{L(\omega,\mu_1)}{L(\omega,\mu_2)}.$$

It is not clear how to estimate μ_1 and μ_2 or, if they are estimated, how to test goodness of fit. Note that in the special case where signal 2 is extremely intense, we can assume that the decision time is negligible, in which case F_2 estimates R and so

$$L(\omega,\mu_1) \rightarrow F_1(\omega)/F_2(\omega) \quad \text{as } \mu_2 \rightarrow \infty.$$

A major unsolved problem in this case is to think through the estimation of parameters, which is not quite conventional since the data have been heavily processed before we get to an estimate of L or l.

Organization of the Paper

The remainder of the paper simply applies the above general theory to a variety of psychophysical tasks: discrimination of an increment in intensity, recognition of one of several intensities, magnitude estimation, simple detection, and simple reaction time. We group these designs not by any operational criterion, but by the extreme theoretical distinction of whether the subject is free to use maximum κ permitted by his buffer store or is "forced" to use a $\kappa = 1$ decision rule. The former arises when there is no time pressure and the signals are both of sufficient duration and intensity, what we call "strong" signals, so that K IATs are very likely to occur on each channel during the presentation. In particular, those tasks, such as discrimination, recognition, and magnitude estimation, in which the signals are so intense that there is no problem whatsoever in detecting

them, seem to permit the use of the maximum (see, however, the section entitled "Estimates of J and K" for a strong qualification of this remark). The latter is assumed to arise when there exists an experimentally imposed time pressure, which can take at least two forms: explicit instructions and payoffs to be as fast as possible or the use of weak signals of brief duration. We do not attempt here to study the difficult intermediate category of experiments in which no explicit time pressure exists, but the intensity-duration of the signal is insufficient, on the average, to fill the buffer store during the presentation of the signal.

We take up these topics in the order: strong signals, weak signals, and RT. The reason for this order is increasing analytic complexity. For strong signals, we derive only statements about choice probabilities and do not study the response times. For weak signals, we derive both probability and temporal relations, but the latter are comparatively simple. Finally, in the RT section we focus almost entirely on the complex problem of detecting the onset of a signal when it can occur at any time.

The basic procedure followed in the analysis of each design is the same. First, we state the exact nature of the experimental task. Next, we describe a decision rule which might be suitable. Finally, we derive equations for the densities of the various decision latencies and for the conditional response probabilities from the decision rule and other assumptions of the theory, calculate Fourier transforms of the densities, and indicate (if we know) how to estimate parameters. Wherever possible, we cite relations to well-known results, either theoretical or empirical.

Parameter Estimates

Problems of estimating parameters arise in any theory of this type, but they are especially acute in the present theory for several reasons. First, the parameters often enter the equations as products— $\mu(s)\beta$, JK, etc.—and so they are not completely identifiable. If $JK = 100$, then either $J = 100$ and $K = 1$ or $J = 1$ and

$K = 100$ is an equally good possibility. This ambiguity cannot be reduced if we stay in a single experimental context. Only by studying behavior in a variety of experiments can we arrive at estimates in which we have some confidence. Second, in many cases we do not have completely satisfactory analytic methods for estimating parameters. Partly this results from difficulties in solving equations; partly because we have not worked out the correct model. For example, we are reasonably confident from various data that $J > 1$; however, at present we have formulated only a $J = 1$ model for simple RTs with random foreperiods. Estimates from this model differ from $J > 1$ models for other experiments by as much as an order of magnitude. Finally, at present we are at the mercy of a few observations from different experiments run in a number of different laboratories for various reasons unrelated to our theory. Nevertheless, to help clarify the theory and to indicate how one might proceed, we make tentative estimates of the basic parameters from existing data. The reader should not take these estimates too seriously; we do not.

Comparison with Other Theories

As we have already suggested, the present theory has much in common with well-known theories; but it differs in significant ways. It is well to appreciate the similarities and differences.

Just as in TSD, we assume that the subject observes a sensory RV and compares it with a criterion or with another sensory RV. Here, however, the RV is one or more IATs, whereas in TSD it is a likelihood ratio, a rather sophisticated function of the sensory datum and the context of the experiment. Since our sensory process is assumed to be Poisson, the sum of $k = \kappa J$ IATs has a gamma distribution with k degrees of freedom, and so the estimate of μ, the reciprocal of the mean of k IATs, has a known distribution different from the normal generally postulated in TSD. Of course, the IATs are temporal RVs and permit us to derive latency distributions, whereas the likelihood ratios have nothing

explicit to do with time. John (1967) has sketched a theory in which an RV, monotonic with time, is compared with a criterion.

Other theories, including TSD, the Poisson counting models, and some simple RT models in which there has been an explicit attempt to deal with time (Carterette, 1966; Edwards, 1965; Estes, 1960; LaBerge, 1962; Laming, 1968; Stone, 1960) differ from ours by assuming the existence of a temporal structure other than that created by the output of the sensory process. The main idea is that there is some sort of time quantum, either imposed by the experimental design (TSD) or by the subject, and within each quantum an RV is observed and compared with a criterion. In most such models, observations are taken until one, or the average, RV gets outside an interval of indecision, that is, a Wald-type sequential test is employed. Unfortunately, it is usually impossible to derive exact closed-form expressions for the resulting latency distributions. By contrast, we postulate no internal temporal structure other than that provided by the output of the sensory process, and the resulting IATs are our sensory RVs; often, latency distributions are easily derived. Of course, if there really is an inherent human time quantum over which information is integrated, as Kristofferson (1967a, 1967b) has argued, it is important to model it correctly; however, the current concensus is against that hypothesis and it seems imperative, therefore, to work out temporal models which do not include any time quantum. Ours do not include such a quantum.

PSYCHOPHYSICS OF STRONG SIGNALS

Assuming that a buffer store exists, we may give a theoretical definition of a strong signal as one of such intensity and duration that, almost certainly, more IATs will occur during its presentation than can be accommodated in the buffer store. More formally, if K is the (IAT) size of the store, T the duration of the signal, and μ its Poisson parameter, then the signal is strong

if μT is considerably larger than K. For strong signals that are also intense (large μ), detection is not the issue; rather, experiments are aimed at understanding the apparent magnitude of the signals or the discriminability of two such signals when their intensities are about equal.

The models we shall investigate, called *limited memory timing*, assume that the subject is allowed, both by the signal and by the reward structure, to accumulate as many IATs as are permitted by his buffer store, namely $k = KJ$, where J is the number of distinct channels. On the average, this requires K/μ seconds per channel. We assume that he uses the average duration of the IATs as an estimate of $1/\mu$, and we derive some probability predictions of this model for discrimination, recognition, and magnitude estimation designs. We do not write down any of the equations for the densities of decision latencies because, with large μ, these latencies are such a small fraction of the response time that they probably cannot be used effectively to study decision rules. (If the present theory is conceptually correct, most simple RT studies, with their very intense signals, elicit information mainly about the residual latency, not the decision one.)

At least two other plausible models exist for strong signals. In the *counting models*, which have been investigated by others, no memory limitation is postulated, and the subject is assumed to count the number of pulses that occur within the signal interval. Just how that interval is demarked for the subject is a moot point. In addition, the counting model has the unfortunate feature of predicting error-free performance as $\mu T \rightarrow \infty$; it is well known that increasing T beyond, at most, a few seconds does not improve performance.

Perhaps the most appealing models from an intuitive point of view are those we may call *sequential, limited memory timing*. The subject is assumed to accumulate information until either the evidence in favor of one response becomes sufficiently persuasive or until the memory is saturated, whichever comes first. Their intuitive appeal is strongly offset by their mathematical complexity. It is often impossible to calculate the distribution of the number of IATs to a decision without making grossly simplifying assumptions.

Discrimination of Intensity

In this section we shall consider pair-comparison designs in which two signals, s_1 and s_2, differing only in intensity, are each presented for duration T on each trial. We let $\mu_i = \mu(s_i)$. The subject is to report which, the first or second, is the more intense in his judgment. The design in which one signal is held fixed and the other varied is called the method of constant stimuli.

We simply assume that as each signal is presented, the subject stores the first $k = KJ$ IATs, records the total time (i.e., forms their sum), compares the two totals, and reports whichever is smaller as the more intense signal. If we let $\mathbf{T}_i$ denote the total time RV arising from signal s_i, then

$$p(1|s_1,s_2) = P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0)$$
$$= P(\mathbf{T}_2/\mathbf{T}_1 \geq 1)$$

is the probability that the first signal is reported as more intense. It is well known (see Parzen, 1962, p. 137) that

$$p(1|s_1,s_2)$$
$$= \frac{(2k-1)!}{[(k-1)!]^2} \int_0^\rho [x(1-x)]^{k-1}dx, \quad [5]$$

where

$$\rho = \frac{\mu_1}{\mu_1 + \mu_2}.$$

This is the incomplete Beta distribution $I_\rho(k,k)$ or, equally, the F distribution

$$Q\left(\frac{1-\rho}{\rho}; 2k,2k\right) = Q\left(\frac{\mu_2}{\mu_1}; 2k,2k\right).$$

It can be written in another way,

$$p(1|s_1,s_2) = \sum_{i=0}^{k-1} f(i; k,\rho)$$

where

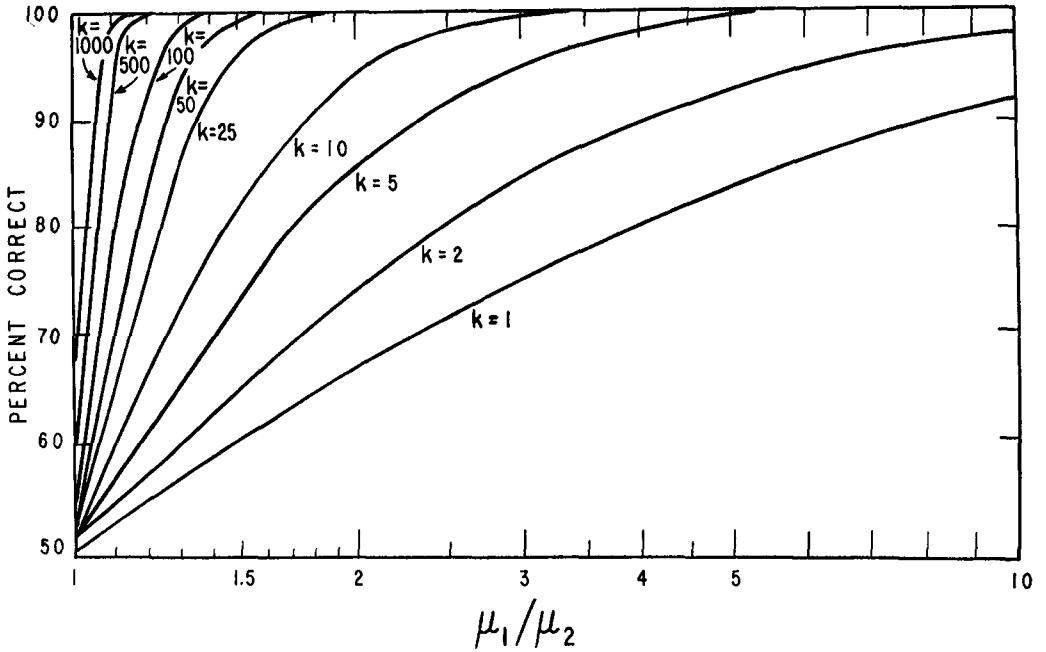$$f(i; k,\rho) = \rho^k \binom{k+i-1}{i}(1-\rho)^i$$

FIG. 2. Variation of percentage of correct response in an intensity discrimination task. (The abscissa is the ratio of the rate parameters for the two signals. The parameter is the number of independent channels on which the decision is based, $k = KJ$.)

is the negative binomial distribution. As Green and McGill (1970) note in a similar model, for large k, Equation 5 is well approximated by

$$p(1|s_1,s_2) \cong \int_{-\infty}^{d'/\sqrt{2}} \mathfrak{N}(0,1), \qquad [6]$$

where

$$d' = \sqrt{k - 1/2}\left(\frac{\mu_1}{\mu_2} - 1\right). \qquad [7]$$

Figure 2 shows $p(1|s_1,s_2)$ as a function of μ_1/μ_2 for several values of k.

Observe that $p(1|s_1,s_2)$ is a constant if and only if μ_1/μ_2 is constant. Since, by hypothesis,

$$\frac{\mu_1}{\mu_2} = \left[\frac{I(s_1) - I_0}{I(s_2) - I_0}\right]^\gamma$$

$$= \left[\frac{I(s_1) - I(s_2)}{I(s_2) - I_0} + 1\right]^\gamma,$$

μ_1/μ_2 is a constant if and only if

$$\frac{I(s_1) - I(s_2)}{I(s_2) - I_0}$$

is a constant, so Weber's law holds for $I(s_2) \gg I_0$. As $I(s_2)$ approaches I_0, two effects contribute to the observed increase in

$$\frac{\Delta I}{I} = \frac{I(s_1) - I(s_2)}{I(s_2)}.$$

Most obviously, I_0 can no longer be neglected and $\Delta I/(I - I_0)$ is constant only if $\Delta I/I$ is increasing. More subtly, as μ is decreased, the number of IATs that can be observed before the signal is terminated must, necessarily, decrease. In effect, the size K of the storage is functionally decreased and, as we show in Appendix A, as long as

$$\frac{\mu_1}{\mu_2} > \frac{1 + 1/(2k - 1)^{\frac{1}{2}}}{1 - 1/(2k - 1)^{\frac{1}{2}}}$$

then in order to maintain $p(1|s_1,s_2)$ at a constant value, μ_1/μ_2 must be increased as k is decreased.

Another interesting feature of the model shows up if, instead of memory limitations, we assume extreme time pressure ($\kappa = 1$) (due to short signal duration) and a single

channel $(J = 1)$ and so set $k = 1$ in Equation 5,

$$p(1|s_1,s_2) = \rho = \frac{\mu_1}{\mu_1 + \mu_2}.$$

Notice that this generalizes nicely to m signals:

$$p(1|s_1,\cdots,s_m) = \int_0^\infty \mu_1 e^{-\mu_1 t} \prod_{i=2}^m e^{-\mu_i t} dt$$

$$= \int_0^\infty \mu_1 e^{-\sum_{i=1}^m \mu_i t} dt$$

$$= \frac{\mu_1}{\sum_{i=1}^m \mu_i}. \qquad [8]$$

Models based on Equation 8 have been called choice models and were investigated by Luce (1959) and others. Within the present framework, we would only expect them to hold for intensity discrimination when the subject is under severe time pressure and has only one channel.

Biased Discrimination of Intensity

If the payoffs are asymmetric, so that it matters to the subject which type of error he makes, then we anticipate biased responses. If his available information is the two random variables T_1 and T_2, the bias could take various forms. One possibility is to say signal 1 when $T_1 - T_2 \geq \epsilon$; another is to say signal 1 when $T_2/T_1 \geq \delta$. With $\epsilon = 0$ and $\delta = 1$, these are equivalent; in general, however, they are distinct. We consider them separately. For the first, the argument given in Appendix A yields

$$P(T_2 - T_1 \geq \epsilon) = \begin{cases} \rho^k \left[\sum_{i=0}^{k-1} \binom{k+i-1}{i} (1-\rho)^i \sum_{j=0}^{k-1-i} \frac{(\mu_2\epsilon)^j}{j!} \right] e^{-\mu_2\epsilon}, & \epsilon \geq 0 \\ (1-\rho)^k \sum_{i=0}^{k-1} \binom{k+i-1}{i} \rho^i \left[1 - e^{\mu_1\epsilon} \sum_{j=0}^{k-1-i} \frac{(-\mu_1\epsilon)^j}{j!} \right] \\ \qquad\qquad + P(T_2 - T_1 \geq 0), \ \epsilon < 0 \end{cases} \qquad [9]$$

where $P(T_2 - T_1 \geq 0)$ is given by Equation 5. Observe that for $k = 1$, Equation 9 reduces to

$$p(1|s_1,s_2) = \begin{cases} \rho e^{-\mu_2\epsilon}, & \epsilon \geq 0 \\ (1-\rho)(1 - e^{-\mu_1\epsilon}) + \rho, & \epsilon < 0 \end{cases} \qquad [10]$$

from which the following receiving operator characteristic (ROC) curve derives:

$$p(1|s_1,s_2) = \begin{cases} \rho\left[\dfrac{p(1|s_2,s_1)}{1-\rho} \right]^{\mu_2/\mu_1}, & p(1|s_2,s_1) \leq 1 - \rho \\ 1 - (1-\rho)\left[\dfrac{1 - p(1|s_2,s_1)}{\rho} \right]^{\mu_2/\mu_1}, & p(1|s_2,s_1) > 1 - \rho \end{cases} \qquad [11]$$

Plots of the ROC curve generated by Equation 9 for several values of k are shown in Figure 3.

The other possibility yields (Appendix A)

$$P\left(\frac{T_2}{T_1} \geq \delta\right) = \frac{(2k-1)!}{[(k-1)!]^2}$$

$$\times \int_0^{\rho(\delta)} [x(1-x)]^{k-1} dx \qquad [12]$$

where

$$\rho(\delta) = \frac{\mu_1}{\mu_1 + \delta\mu_2}.$$

For $k = 1$,

$$p(1|s_1,s_2) = \frac{\mu_1}{\mu_1 + \delta\mu_2}$$

from which we obtain the ROC curve

$$\frac{p(1|s_1,s_2)}{p(2|s_1,s_2)} = \left(\frac{\mu_1}{\mu_2}\right)^2 \frac{p(1|s_2,s_1)}{p(2|s_2,s_1)},$$
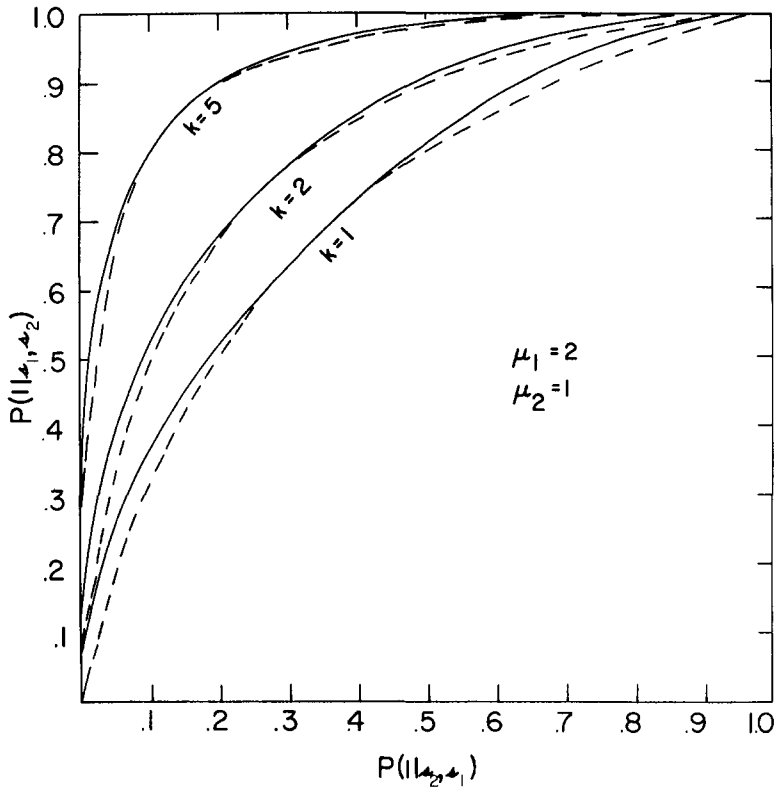
FIG. 3. ROC curves for intensity discrimination experiment with $\mu_1 = 2$, $\mu_2 = 1$ and two different decision rules. (The solid curves are generated by varying the criterion on the *difference* on the sum of k IATs associated with each interval. The dotted curves are generated by varying the criterion on the *ratio* of the sum of k IATs associated with each interval.)

which is identical to that of the choice model (see Luce, 1963a). The curves generated from Equation 12 for various values of k are also shown in Figure 3. It is unlikely that experimental data could be used to distinguish between the curves based on Equations 9 and 12.

Before leaving our discussion of the discrimination of intensity differences, we should comment on an obvious deficiency of models in this area. The present timing model says nothing about the interval separating the two signal presentations. According to the model, it is irrelevant, whereas, in reality, discriminability does diminish with increasing separation between the two signals. The obvious way to remedy this situation within the counting model is to assume that the count is not recalled with perfect fidelity (Kinchla & Smyzer, 1967).

Similar assumptions might be made within the timing model concerning the fidelity of the time recorded for the first signal. Various alternatives are possible, but none has been worked through to any interesting predictions; we leave this as a conspicuous open problem.

Recognition of One of m Signals

By a recognition (absolute identification) design we mean that on each trial, one of m signals, $s_1, \cdots, s_m$, is presented and the subject is to identify which by using one of m responses that are in one-to-one correspondence with the signals. Usually the correspondence is established as being order preserving. In our analysis we assume the signals differ only in intensity, hence s_{i+1} is more intense than s_i for all i. Our approach to recognition parallels that

of Thurstonian psychophysics, except we use a gamma distribution rather than the normal. In particular, we assume that the signal duration is sufficiently long and the intensities sufficiently great that, almost certainly, k IATs can be observed, and that their sum is a random variable $\mathbf{T}_i$ which has a gamma density with parameters k and μ_i when signal s_i is presented. We assume that the subject selects $m - 1$ criterion values

$$0 < \beta_{m-1} < \cdots < \beta_p < \beta_{p-1} < \cdots < \beta_1 < \infty,$$

and that his decision rule is to respond p if and only if

$$\beta_p \leq \mathbf{T} < \beta_{p-1}.$$

Clearly, then

$$p(\mathbf{R} \leq p\,|\,s_i) = \int_{\beta_p}^{\infty} \frac{\mu_i{}^k t^{k-1} e^{-\mu_i t}}{(k-1)!}\,dt$$
$$= \int_{\beta_p \mu_i}^{\infty} \frac{y^{k-1} e^{-y}}{(k-1)!}\,dy. \quad [13]$$

For large k the central limit theorem yields

$$P(\mathbf{R} \leq p\,|\,s_i) \cong \int_{-\infty}^{z_i(p)} \mathfrak{N}(0,1), \quad [14]$$

where

$$z_i(p) = \sqrt{k} - \frac{\mu_i \beta_p}{\sqrt{k}}. \quad [15]$$

Thus, Equation 13 is approximately a non-equal variance Thurstonian model. The equal variance case, in which $z_i(p) = (\beta_p - \mu_i)/\sigma$, has been proposed for intensity recognition by Durlach and Braida (1969) and is tested experimentally in Braida and Durlach (1972).

Given recognition data, the $(m-1)m$ $z_i(p)$ can be obtained from Equation 14 (the cases where $z_i(p) = -\infty$ or ∞ are omitted from the estimation scheme). We can then use Equation 15 in the following (non-optimal) way to estimate parameters. We note that from Equation 15 it follows immediately that

$$z_i(p) - z_j(p) = (\mu_j - \mu_i)\beta_p/\sqrt{k}$$

and

$$z_i(p) - z_i(q) = \mu_i(\beta_p - \beta_q)/\sqrt{k}.$$

Thus, if the model is correct, the geometric mean of

$$\frac{z_i(p) - z_j(p)}{z_i(q) - z_j(q)} = \frac{\beta_p}{\beta_q}$$

over all i, j, $i > j$ yields an estimate of β_p/β_q. Similarly, the geometric mean of

$$\frac{z_i(p) - z_i(q)}{z_j(p) - z_j(q)} = \frac{\mu_i}{\mu_j}$$

over all p, q, $p > q$ yields an estimate of μ_i/μ_j.

Since $\mu_i > 0$ and the unit of measurement is not determined, there is no loss of generality in choosing the unit so that

$$\sum_{i=1}^{m} \mu_i = 1.$$

We then have as an estimate of μ_i,

$$\hat{\mu}_i = \frac{1}{m} \sum_{j=1}^{m} \left(\frac{\mu_i}{\hat{\mu}_j}\right).$$

With that in hand, we can use Equation 15 to estimate $\beta_p/\sqrt{k}$ by forming the geometric mean of

$$\frac{z_j(p) - z_i(p)}{\hat{\mu}_i - \hat{\mu}_j}$$

over all i, j, $i > j$. Finally, from Equation 15 we see that the mean over all i and p of

$$\hat{\mu}_i\left(\frac{\beta_p}{\sqrt{k}}\right) + z_i(p)$$

is an estimate of $\sqrt{k}$.

With all of these parameters estimated, we may use Equation 15 to predict the $m(m-1)$ (less the ∞ cases) observed $z_i(p)$ from the $2m - 1$ free parameters. These predictions can be displayed as a scatter diagram and evaluated by a chi-square test.

It should be of interest to compare the estimate of μ_i/μ_j obtained in this way with those obtained by magnitude estimation (see the next section).

Magnitude Estimation

Experimentally, magnitude estimation (ME) and recognition have much in common: a single signal is presented on each trial, and a numerical response is required in which a larger number is assigned to a

more intense signal. They differ, of course, in that the experimenter establishes a one-to-one correspondence in the recognition design between the responses "1," "2," $\cdots$, "m" and a fixed set of m signals; whereas he asks the subject in the ME design to reveal his "built-in" one-to-one correspondence between all positive real numbers and all possible signals for an unspecified sample of these signals. One notable consequence of this difference is that there are correct and incorrect responses in the recognition experiment, and so payoffs can be used, whereas no concept of error is appropriate in the ME design. Our theoretical analysis differs considerably.

We simply assume that the subject attempts to estimate $c\mu$, where c is an adjustable unit parameter, by observing k IATs and using the total time **T** as his information about the signal. In fact, we assume that he responds $\mathbf{M} = ck/\mathbf{T}$ as his estimate of the magnitude of the signal. Since **T** is distributed as a kth order gamma, its reciprocal has the Pearson Type V distribution (see Kendall & Stuart, 1958):

$$P(\mathbf{M} = x) = \frac{(kc\mu)^k e^{-kc\mu/x}}{(k-1)! x^{k+1}}. \quad [16]$$

Perhaps the most unusual feature of this density is its high tail—of the form $x^{-(k+1)}$ as x becomes large. This agrees with the observations of Luce and Mo (1965) and Schneider and Lane (1963). In fact, the tails are so high that moments beyond $k - 1$ do not exist. The mode, mean (for $k > 1$), geometric mean (for $k > 1$), and variance (for $k > 2$), which are easily derived, are given by:

$$\left(c\frac{k}{k+1}\right)\mu, \quad \left(c\frac{k}{k-1}\right)\mu, \quad (cke^{C - \sum_{i=1}^{k+1} 1/i})\mu,$$

$$\left(c\frac{k}{k-1}\right)^2 \frac{1}{(k-2)}\mu^2, \quad [17]$$

respectively, where C is Euler's constant ($\cong .577216$). Since the observed mean or median magnitude estimates are, approximately, power functions of signal energy (S. S. Stevens, 1957), the postulate that μ is a power function of intensity is reasonable. Moreover, this gives a way to estimate the

exponent γ. This assumption was used to account for Weber's law.

If this model is correct, it is clearly better not to work with the data and Equation 16 directly, but rather to form reciprocals. These have the density

$$P\left(\frac{1}{\mathbf{M}} = y\right) = \frac{(ck\mu)^k y^{k-1} e^{-ck\mu y}}{(k-1)!} \quad [18]$$

which, because of its exponential tail, has much more stable moments than Equation 16. The mean and variance both exist and are

$$\frac{1}{c\mu} \quad \text{and} \quad \frac{1}{(c\mu)^2 k},$$

and so we have the following estimate of k:

$$\hat{k} = \text{mean}^2/\text{variance}. \quad [19]$$

If all forms of signal intensity are recoded in pulse trains with a rate that increases with intensity, then comparisons of intensity between modalities is a possibility. At least intensities on different modalities are represented by a common measure, time, and so we need only postulate that the nervous system is capable of making comparisons of IATs originating in the different senses. Thus, the a priori mysterious, but successful, method of cross-modality matching makes complete sense in this theory. Because of our power function assumption, we predict a power function matching relation with an exponent that is the ratio of the two ME exponents, which is what is observed. In addition, since both signals exhibit variability in their IAT representation, we predict that the distribution of matches should be more variable than the reciprocal of the corresponding MEs. We do not know of data bearing on this prediction.

If we approach ME data much as we did recognition data, we have from Equation 16

$$P(\mathbf{M} \leq p \,|\, s_i) = \int_0^p \frac{(kc\mu_i)^k e^{-kc\mu_i/x}}{(k-1)! x^{k+1}} dx$$

$$= \int_{kc\mu_i/p}^\infty y^{k-1} e^{-y} dy / (k-1)!$$

$$\cong \int_{-\infty}^{z_i(p)} \mathfrak{N}(0,1), \quad [20]$$

where

$$z_i(p) = \sqrt{k}\left(1 - \frac{c\mu_i}{p}\right). \qquad [21]$$

Here p is a cut point established by us, not by the subject. Considering only the μ_i term, the dependence on k and the criterion p are inverse to that found in the recognition model (Equation 15).

Estimates of J and K

According to the ME model (especially Equation 17) and S. S. Stevens' (1957) data, we know that for 1000-Hz. tones, approximately,

$$\frac{\mu_1}{\mu_2} = \left(\frac{I_1}{I_2}\right)^{.3},$$

where I is acoustic power. It is also well known that the Weber fraction corresponding to $p(1|s_1,s_2) = .75$ is about .08, so the corresponding parameter ratio is

$$\frac{\mu_1}{\mu_2} = (1.08)^{.3} = 1.023.$$

Using Equation 6, the normal approximation to Equation 5, we have an estimate of $k = JK$ of about 1750.

With k known, it is sufficient to estimate either J or K. The only methods that have occurred to us are for J. One possibility, using ME data, is suggested by data of S. S. Stevens (1966) and J. C. Stevens and Hall (1966), which were both concerned with the dependence of MEs on signal duration T. For both brightness and loudness, the MEs increase linearly in log T and then abruptly become constant. In the case of brightness, the location of the corner is a function of intensity, being at about 10 msec. for an 85-decibel (db.) signal and at about 100 msec. for a 55-db. one. This strongly suggests that the subject does not fill his buffer store before determining his estimate, but rather that only a few IATs are collected on each channel. The times indicate a 10 to 1 ratio in the μs over a 30-db. range which agrees with the fact that the range is equivalent to a stimulus ratio of 10^3 and that the brightness ex-

ponent is about 1/3. Indeed, the numbers are such that on the basis of threshold data, using fixed intervals, we are probably safe in assuming that only one IAT is obtained from each channel.

The situation for loudness differs in one important respect: the corner, which is at 150 msec., is independent of intensity. This strongly suggests that the number of IATs accepted increases with intensity, which agrees qualitatively with the finding mentioned by Luce and Mo (1965) that their estimate of k increases; however, they reported a factor of only 4 over a 47-db. range whereas we would predict, using an exponent of .3, a factor of 26 in order for the location of the corner to remain fixed. In contrast Schneider and Lane (1963) using magnitude production, but converting their data to an equivalent estimation form, found the standard deviation of the estimates to be 1/5 of the mean. From Equation 17, this yields the estimate of $k = 27$.

There appears to be a discrepancy here in that the Stevens and Hall data for loudness estimation suggest that k increases with intensity, whereas those of Schneider and Lane for loudness production suggest it does not. Further investigation is required to determine if this is a difference between estimation and production or if something deeper is wrong with the theory.

Assuming only one IAT per channel, $\kappa = 1$, for the Schneider and Lane experiment, $J = 27$ and so $K = 1750/27 \cong 65$.

The reader is no doubt perplexed, as are we, about why 65 IATs should be collected per channel in a discrimination experiment and only 1 or a few in the ME procedure. The only explanation that we have thought of, and it is wholly ad hoc, is that calculating a reciprocal is considerably more complex than comparing two sums, and so vast amounts of storage capacity are reserved to carry it out. We would be more convinced of this had the loudness data been similar to the brightness data and so suggested $\kappa = 1$, independent of intensity. There seems to be a problem in need of better explanation.

Psychophysics of Weak Signals

We turn next to one of the classical areas of psychophysics, the detection of weak signals. Here the focus continues to be on various response probabilities, but in contrast to strong signals the response times begin to assume some interest because the delays of the sensory-decision process are no longer negligible relative to delays of the residual process.

The essential problem for the subject is to decide on each trial whether or not a signal is present in a noise background. Because the signal is weak, we know that the density of Poisson events is low or, to put it another way, that the expected value of an IAT is moderately large. We may impose time pressure on the subject either by deadlines and payoffs or by instructions while leaving the signal on until a response occurs, as in a type of RT experiment with catch trials, or by using signals of such short duration that a failure to decide quickly almost certainly brings into play pulses not due to the signal, as in a conventional yes–no design. We shall suppose that both of these procedures invoke decision rules that involve only a single or very few IATs from each channel.

It is not customary to regard fixed-interval detection experiments as involving any strong time pressure, but clearly they do if anything like the present theory is correct. For example, suppose that a 500-msec. interval is specified and that a faint signal is presented with, let us say, a value of $\mu = 4$, then the expected time to the first event is 250 msec. and the second is 500 msec. This suggests that on the average, it would be futile to wait for a second IAT and that, indeed, it may be futile to wait for even one if the second impulse lies beyond the end of the signal interval.

We examine situations having only two signals: noise of constant power, n, and a signal of fixed frequency and intensity in that noise, s. Thus, there are only two Poisson parameters, $\mu = \mu(s)$ and $\nu = \mu(n)$, where $\mu > \nu$. We use obvious mnemonics for the responses: Y for yes, N for no, 1 for the first interval, 2 for the second. A typical conditional response probability is denoted $p(Y \mid n)$, that is, the probability of a yes response given that no signal was presented. The response time and decision latency densities are subscripted similarly, so for example, f_{nY} is the density of the yes response times to the presentation of no signal, and l_{sN} is the density of the no decision latencies to the presentation of a signal.

Yes–No Design with Response-Terminated Signals

It is convenient to begin with an unconventional yes–no design because it is one of the simplest to analyze. With that as an example, we turn later to the usual fixed-interval design. Consider an experiment in which the potential onset of a signal is clearly indicated to the subject (by an intense signal in another modality, e.g.). On some trials there is no signal. On the others, a constant intensity signal is presented and remains on until the subject responds Y or N. The subject is pressed (by payoffs or instructions) to respond as fast as he can. In other words, this is a simple RT experiment with zero foreperiod and catch trials, but with a faint response-terminated signal (which is not usual in the RT literature).

The decision rule we shall suppose is that the first κ IATs are obtained from each channel and if the average of these $k = \kappa J$ IATs is less than β, respond Y; otherwise respond N. The decision time is then simply the time of the slowest $(\kappa + 1)$ pulse, which has the $(\kappa + 1)$-order gamma distribution $\gamma_{\kappa+1,\mu}(t)$ with intensity parameter μ. Since if J RVs have a common density, that of the largest is

$$l_s(t) = \frac{d}{dt}\left[\int_0^t \gamma_{\kappa+1,\mu}(x)dx\right]^J$$

$$= J\gamma_{\kappa+1,\mu}(t)\left[\int_0^t \gamma_{\kappa+1,\mu}(x)dx\right]^{J-1}. \quad [22]$$

From this, a direct calculation yields the mean reaction time (MRT),

$$\mathrm{MRT}_s = \bar{r}_Y p(Y \mid s) + \bar{r}_N P(N \mid s)$$

$$+ \frac{h(J,\kappa)}{\mu}, \quad [23]$$

where

$$h(J,\kappa) = J \int_0^\infty \frac{x^{\kappa+1} e^{-x}}{\kappa!}$$
$$\times \left[\int_0^x \frac{y^\kappa e^{-y} dy}{\kappa!} \right]^{J-1} dx \quad [24]$$

and $\bar{r}_Y$ and $\bar{r}_N$ are the mean residual latencies for, respectively, Y and N responses. For large κ, Equation 24 can be transformed by the central limit theorem into

$$h(J,\kappa) \cong \kappa + 1 + \sqrt{\kappa + 1} H(J),$$

where

$$H(J) = J \int_{-\infty}^\infty x \mathfrak{N}(0,1) \left[\int_{-\infty}^x \mathfrak{N}(0,1) \right]^{J-1} dx$$

is the mean of the largest of J normally distributed RVs with mean 0 and variance 1. Several values of $H(J)$ are given in Table 1.

To get some idea of the error of the approximation for small κ, we compute $h(2,1) = 2.750$ as compared with the approximate value of 2.797, an error of less than 2%.

Note that if $\bar{r}_Y = \bar{r}_N$, then if we can vary κ experimentally, Equation 23 implies

$$\text{MRT}_n = \frac{\mu}{\nu} \text{MRT}_s + \bar{r} \left(1 - \frac{\mu}{\nu} \right). \quad [25]$$

This linear regression provides a test of the model as well as a way to estimate μ/ν and $\bar{r}$.

Although the density of decision times is not simple in this model, the expression for the response probabilities is. Let $\mathbf{T}$ denote the sum of $k = J\kappa$ IATs. Because each IAT is exponentially distributed, $\mathbf{T}$ has a gamma distribution of order k. Thus,

$$p(Y|s) = P(\mathbf{T} \le k\beta)$$
$$= P \left(\frac{1}{\mathbf{T}} \ge \frac{1}{k\beta} \right)$$
$$= \int_{\frac{1}{k\beta}}^\infty g_s(x) dx, \quad [26]$$

where

$$g_s(x) = \frac{\mu^k e^{-\mu/x}}{(k-1)! x^{k+1}}, \quad [27]$$

TABLE 1

$H(J)$ versus J

J	$H(J)$
2	.564
5	1.163
10	1.539
20	1.867
60	2.319
100	2.508
200	2.746
500	3.037
1000	3.241

Note.—These values are taken from Table I of Tippett (1925). Values of $H(J)$ for all values of J, $2 \le J \le 1000$, can be obtained by taking $\frac{1}{2}$ of the mean range values given in Table X of Tippett (1925) or, for less accurate work, they may be interpolated on a plot of $H(J)$ versus $\log J$, which is nearly linear.

which is a Pearson Type V density (see Kendall & Stuart, 1958).

As we have seen in Equation 17, the mean m of g_s exists only for $k > 1$ and the variance of σ^2 only for $k > 2$, and

$$m = \left(\frac{k}{k-1} \right) \mu,$$
$$\sigma^2 = \left(\frac{k}{k-1} \right)^2 \left(\frac{k}{k-2} \right) \mu^2.$$

Thus, for $k > 2$, the difference between the signal and noise densities normalized by the standard deviation of the latter is

$$d' = \sqrt{k-2} \left(\frac{\mu}{\nu} - 1 \right). \quad [28]$$

If we approach the problem as in TSD, we consider the likelihood ratio

$$L(x) = \left(\frac{\mu}{\nu} \right)^k e^{-(\mu-\nu)x},$$

and so

$$lnL(x) \ge \eta \text{ iff } x \le c = \frac{kln\frac{\mu}{\nu} - \eta}{\mu - \nu},$$

which, therefore, leads to exactly the same equation as 26, with c replacing $k\beta$.

By substituting $z = \mu/x$ in Equation 26,

$$p(Y|s) = \int_0^{\mu k\beta} \frac{z^{k-1} e^{-z}}{(k-1)!} dz, \quad [29]$$

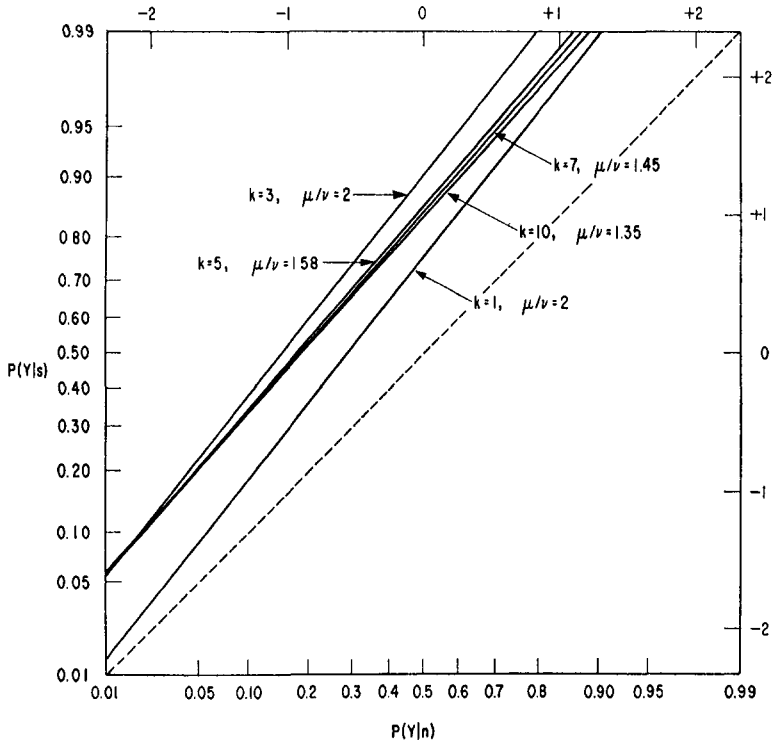which is the incomplete gamma function.

Fig. 4. ROC curve for first κ IATs from J parallel channels. (The parameter k is the product of $J\kappa$. The parameter μ/ν is the ratio of the rate parameter associated with the signal compared to the rate parameter for noise alone.)

For large k, the gamma density approaches the normal $\mathfrak{N}(k,\sqrt{k})$ and so, approximately,

$$p(Y|s) \cong \int_{-\infty}^{\sqrt{k}(\mu\beta-1)} \mathfrak{N}(0,1). \qquad [30]$$

A similar expression holds for $p(Y|n)$. Let $z(s)$ and $z(n)$ denote the normal deviates corresponding to $p(Y|s)$ and $p(Y|n)$,

$$z(s) = \frac{\mu}{\nu}z(n) + \sqrt{k}\left(\frac{\mu}{\nu} - 1\right). \qquad [31]$$

Thus, on normal–normal paper, the ROC curve is approximately a straight line with slope μ/ν. Note that by setting $z(n) = 0$,

$$d' = \sqrt{k}\left(\frac{\mu}{\nu} - 1\right), \qquad [32]$$

which differs from the true value by replacing $k - 2$ by k. For measurable levels

of detectability ($d' \approx 1$) the slope μ/ν is approximately $1 + 1/\sqrt{k}$ which approaches 1 as k increases. As Figure 4 shows, the normal approximation is excellent for all k; however, the actual slopes are considerably less than μ/ν for k as large as 10. This results from errors in approximating the gamma distribution by the normal distribution.

It should be noted that the shape of the ROC curve is very unusual. For example, with $k = 1$ it is

$$p(Y|s) = 1 - [1 - p(Y|n)]^{\mu/\nu},$$

which is the power function ROC curve shown in Figure 5. This is different from the power relation,

$$p(Y|s) = p(Y|n)^{\alpha},$$

suggested by Egan, Greenberg, and Schulman (1961).

Some caution is required in interpreting this prediction since most published data are for fixed-interval designs and, as is shown later (Equation 35), the ROC curve for the corresponding $J = 1$ fixed-interval case is approximately of the correct shape. Moreover, as we have shown empirically (Green & Luce, 1972), even with response-terminated signals, what happens depends on the payoff structure. With an imposed deadline such that responses later than the deadline are penalized, the slope of the ROC curve in a normal probability plot is 1 or slightly less. However, if the deadline is imposed only on signal trials, the slopes are considerably greater than 1.

Speed–Accuracy Tradeoff

At least in situations with response-terminated signals, it has been generally felt that within limits, subjects should be able to increase their accuracy of performance at the expense of taking more time. This is not a matter of variation over the ROC curve, in which case a change in criterion decreases one of the errors, for example, $p(Y|n)$, only at the expense of increasing the other, $p(N|s)$. In a speed–accuracy tradeoff, both types of errors are to be reduced at the expense of taking more time to respond. Empirically, the results have been ambiguous, ranging from fast guess results in which the subject completely fails to identify which signal is presented or identifies it nearly perfectly (Swensson, 1969, Experiment I; Swensson & Edwards, 1971) to difficult-to-discriminate signals in which he exhibits a tradeoff (Swensson, 1969, Experiments II and III).

Within the present model, such a tradeoff can be effected only by using more than one IAT per channel on which to base the response. Obviously, increasing the number of IATs on which to base an estimate of $1/\mu$ decreases the variability of the estimate and takes more time. In fact, from the approximation to $h(J,\kappa)$ above, we see from Equation 23 that

$$\text{MRT}_s(\kappa + 1) - \text{MRT}_s(\kappa)$$

$$\cong \frac{1 + [(\kappa + 2)^{\frac{1}{2}} - (\kappa + 1)^{\frac{1}{2}}]H(J)}{\mu} \geq \frac{1}{\mu}$$
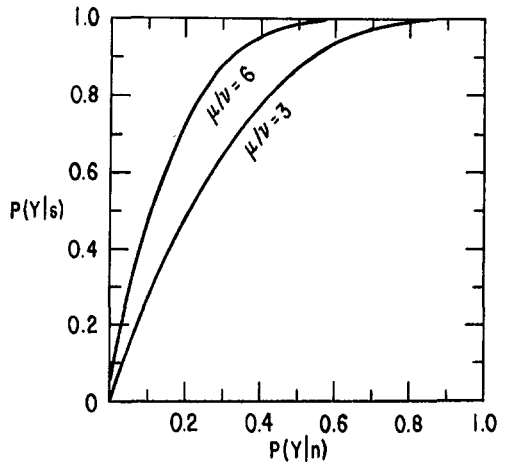

FIG. 5. The $k = 1$ ROC curve for response terminated yes–no design. (The ordinate is the probability of a Y response given that a signal was presented. The abscissa is the probability of a Y response given that no signal was presented. The parameter is the ratio of the rate parameter associated with the signal compared to the rate parameter for noise alone.)

Thus, if the signals are very intense and so μ is large, a negligible change occurs in the response time; whereas, if the signals are very faint and μ is small, a large change occurs and the tradeoff appears to be discontinuous. For example, with $\mu = 4$, the change is at least 250 msec., which is likely to be interpreted as an observer who can operate only in two modes. For signals of moderate intensity, say μ between 20 and 50, we are likely to see a more "continuous" trade of speed and accuracy because in that range each additional IAT adds, on the average, from 20 to 50 msec.

Yes–No Design with Fixed-Signal Duration ($J = \kappa = 1$ Model)

The problem of how to deal with interval markers becomes perplexing when we analyze the conventional yes–no design. The situation is exactly the same as the preceding one except that the signal is terminated T seconds after its onset, independent of the subject's response; moreover, the period during which it might be present is clearly marked by a strong

signal in another modality. On all trials, we wait for a response to occur, and we record Y to an s presentation as a correct response even if the response occurs long after the termination of the signal.

The major question in analyzing this experimental design is what decision rule to use. The problem is whether or not the nervous system takes into account the information about the end of the signal interval. If it does not and if the first IAT does not lie wholly within the signal interval, then the second pulse is generated by noise, not by signal. If it does, then we have a type of deadline model. We derive the various RT densities and conditional response probabilities for both rules on the assumption $J = 1$, and then show that the second is inconsistent with some RT data. The model for general J is complicated.

Case 1: Decision based on the length of the first IAT independent of when it occurs. This is exactly the same rule as was used above but with $J = 1$; it ignores the information about the end of the signal interval. The equations for n trials are, of course, unchanged (Equation 29 with ν substituted for μ, n for s, and with $k = 1$). The signal case is a good deal more complex because the signal may end before the second pulse arrives. In Appendix B we derive the response density, from which one can derive the following expressions for the response probabilities and MRTs:

$$
p(Y \mid s) =
\begin{cases}
1 - e^{-\mu\beta} - (1 - e^{-\nu\beta})e^{-\mu T} \\
\quad \times \left(\dfrac{\mu}{\nu} - 1\right), \quad \beta \leq T \\[2mm]
1 - e^{-\mu T}e^{-\nu\beta}\left(1 - \dfrac{\mu}{\nu}\right. \\[2mm]
\quad \left. + \dfrac{\mu}{\nu}e^{\nu T}\right), \quad \beta > T
\end{cases}
\quad [33]
$$

$$
p(Y \mid n) = 1 - e^{-\nu\beta}. \qquad [34]
$$

Eliminating β, we have for the ROC curve

$$
p(Y \mid s) =
\begin{cases}
1 - [1 - p(Y \mid n)]^{\mu/\nu} \\
\quad - p(Y \mid n)e^{-\mu T}\left(\dfrac{\mu}{\nu} - 1\right), \\
\qquad p(Y \mid n) \leq 1 - e^{-\nu T} \\[2mm]
1 - [1 - p(Y \mid n)]e^{-\mu T} \\
\quad \times \left(1 - \dfrac{\mu}{\nu} + \dfrac{\mu}{\nu}e^{\nu T}\right), \\
\qquad p(Y \mid n) > 1 - e^{-\nu T}.
\end{cases}
\quad [35]
$$

Observe that the upper limb, $p(Y \mid n) > 1 - e^{-\nu T}$, is linear and the lower, curvilinear; some have felt that the data support a curve of this character. For $\beta \leq T$,

$$
\text{MRT}_{sY} = \bar{r}_Y + \left\{ \dfrac{2}{\mu}(1 - e^{-\mu\beta}) \right.
$$

$$
- \beta e^{-\mu\beta} + \left(\dfrac{\mu}{\nu} - 1\right)e^{-\mu T}
$$

$$
\times \left[\beta(1 + e^{-\nu\beta}) - \left(T + \dfrac{2}{\nu}\right) \right.
$$

$$
\left. \left. \times (1 - e^{-\nu\beta})\right]\right\} \Big/ p(Y \mid s) \quad [36]
$$

$$
\text{MRT}_{sN} = \bar{r}_N + \left\{ \left(\dfrac{1}{\mu}\right) + \beta\,e^{-\mu\beta} \right.
$$

$$
+ \left(\dfrac{\mu - \nu}{\nu}\right)\left[\left(T + \dfrac{1}{\nu} + \dfrac{1}{\mu}\right)\right.
$$

$$
- \left(T + \dfrac{1}{\nu} + \beta\right)e^{-\nu\beta}\right]
$$

$$
\left. \times e^{-\mu T}\right\} \Big/ p(N \mid s). \quad [37]
$$

The calculation of the MRT equations for $\beta > T$ is left as an exercise.

Some Data

Observe that if the false alarm rate is held constant (i.e., ν and β are constant in Equation 34), then MRT_{sY} (Equation 36) decreases with increasing T, as was reported by Grossberg (1968) for brightness.

In fact, if $\mu\beta$ is sufficiently small so that $1 - e^{-\mu\beta} \cong \mu\beta$, then

$$\text{MRT}_{sY} \cong \bar{r}_Y + \begin{cases} \dfrac{1 + \nu\beta - (\mu - \nu)T}{\nu} \\ \qquad \text{if} \quad e^{-\mu T} \cong 1 - \mu T \\ \dfrac{1}{\mu} + \beta, \quad \text{as} \quad T \to \infty. \end{cases}$$

Since Grossberg obtained a nearly negligible false alarm rate, β must have been exceedingly small. So for small T, for example, 10 msec., we see that

$$\text{MRT}_{sY} \cong \bar{r}_Y + \begin{cases} \dfrac{1}{\nu}, & \beta, T \ll 1 \\ \dfrac{1}{\mu}, & \mu T \to \infty. \end{cases}$$

Taking as a reasonable guess $\bar{r}_Y = 100$ msec., his data place ν in the range of 3–4 and for these faint signals, μ in the range of 3.5–7.

We next examine whether Bloch's law obtains in our model. For small $\nu\beta$ and νT, Equation 33 reduces to:

$$p(Y|s) \cong \begin{cases} 1 - e^{-\mu\beta} - (\mu - \nu)\beta e^{-\mu T}, \\ \qquad\qquad\qquad\qquad \beta \leq T \\ 1 - e^{-\mu T}(1 - \nu\beta + \nu T), \\ \qquad\qquad\qquad\qquad \beta > T. \end{cases}$$

If we assume both that $\mu \gg \nu$ and μT are sufficiently small so that terms greater than $(\mu T)^2$ can be neglected, then this becomes

$$p(Y|s) \cong \begin{cases} \beta\mu^2(T - \beta/2), & \beta \leq T \\ \nu\beta(1 - \mu T) + \mu T \\ \qquad\qquad - (\mu T)^2/2, & \beta > T. \end{cases}$$

Holding the false alarm rate, and so $\nu\beta$, constant, we see that $p(Y|s)$ constant implies, for small β, that $\mu^2 T$ is constant, whereas increasing β alters this until, for $\beta > T$, μT is constant.

According to S. S. Stevens (1966), MEs for brief flashes rise as from the .4 to .5 power of I. If the latter, then $\mu^2 T$ constant is equivalent to IT constant, which

is Bloch's law. This is well confirmed for small T. According to the model, for $T < \beta$ we should find that it is replaced by the constancy of $I^{\frac{1}{2}}T$, which is not reported. Turning to the loudness data, Stevens and Hall (1966) report that near threshold, $I^{.8}T$ is a constant for equal loudness, which for the model to hold, requires μ to be the .4 power of I. The usually quoted figure is .3; moreover, J. C. Stevens and Hall found that, unlike the brightness data, the exponent is unaffected by the brevity of the signal. However, it is well known that near threshold the apparent exponent is considerably in excess of .3 because the dependence is on $I - I_0$, not on I. So this model does not seem grossly inconsistent with what is known.

Case II: Decision based only on IATs within $(0,T)$. The decision rule is this. Choose $\beta \leq T$. Respond Y if at least two pulses occur within $(0, T)$ and if the first IAT is less than β; otherwise, respond N. In Appendix B, we show that

$$p(Y|s) = 1 - e^{-\mu\beta} - \mu\beta e^{-\mu T}, \quad \beta \leq T. \quad [38]$$

We cannot get an explicit formula for the ROC curve for $\beta \leq T$ and it is obscure what to assume for $\beta > T$. Numerical calculation yields the former. For the latter, we can either assume the subject stays at the value $\beta = T$, jumps to $(1, 1)$, or forms a linear combination of these two. The latter is shown in Figure 6.

The MRTs (Appendix B) are:

$$\text{MRT}_{sY} = \bar{r}_Y + \left[\frac{2}{\mu}(1 - e^{-\mu\beta}) \right.$$

$$- \beta e^{-\mu\beta} - (\mu T + 1)$$

$$\left. \times \beta e^{-\mu T} \right] \Big/ p(Y|s). \quad [39]$$

$$\text{MRT}_{sN} = \bar{r}_N + \left[\left(\frac{1}{\mu} + \beta \right) e^{-\mu\beta} \right.$$

$$\left. + \left(T\mu\beta - \frac{1}{\mu} \right) e^{-\mu T} \right] \Big/$$

$$p(N|s). \quad [40]$$

This model is almost surely wrong. We note from Equation 39 and $\beta \leq T$, and if $\mu\beta$ is small,

$$\mathrm{MRT}_{,Y} \rightarrow \hat{r}_Y + \begin{cases} \dfrac{\beta}{\mu T} & \text{if } e^{-\mu T} \cong 1 - \mu T \\ \dfrac{1}{\mu} + \beta & \text{as } T \rightarrow \infty. \end{cases}$$

As is easily seen, for $\beta \leq T$, $\dfrac{\beta}{\mu T} < \dfrac{1}{\mu} + \beta$, contrary to Grossberg's (1968) data.

Two-Alternative Forced-Choice Design
(J = κ = 1 Model)

In this well-known design, two listening intervals are defined. They are each of duration T and are separated by a short time. On each trial, the signal is presented in exactly one interval and the subject is required to say which. The possibility exists here, as in the yes–no design, that the subject bases his decision on events outside the listening interval. We assume not, however, and so we assume the natural analogue of the decision rule just studied: the first IAT is observed in each interval and the one having the shorter IAT is reported as containing the signal. The only difficulty, which is serious with weak signals, arises when at least one interval

does not contain any IAT; in that case, we let T be the conventional length of the first IAT. Since this can happen with positive probability in each interval, there is a positive probability of an unresolved decision; we assume that with bias b the subject chooses the first interval.

We will not attempt to calculate the densities of decision times, which of course is possible, but will content ourselves with the response probabilities. Let the RV of the time of arrival of the ith pulse be denoted $\mathbf{Y}_i$. Define the decision random variable

$$\mathbf{X} = \begin{cases} \mathbf{Y}_2 - \mathbf{Y}_1, & \text{if } \mathbf{Y}_2 \leq T \\ T, & \text{if } \mathbf{Y}_2 > T. \end{cases}$$

Thus,

$$P(\mathbf{X} = t \mid s)$$
$$= \begin{cases} \displaystyle\int_0^{T-t} \mu e^{-\mu x} \mu e^{-\mu t} dx, & t < T \\ e^{-\mu T} + \displaystyle\int_0^T \mu e^{-\mu x} e^{-\mu(T-x)} dx, & t = T \end{cases}$$
$$= \begin{cases} \mu(e^{-\mu t} - e^{-\mu T}), & t < T \\ \mu e^{-\mu T}(1 + \mu T), & t = T \end{cases}$$

$P(\mathbf{X} = t \mid n)$ is the same with ν substituted for μ. Assuming the independence of the two intervals,

$$p(1 \mid s,n) = P(\mathbf{X}_1 - \mathbf{X}_2 \leq 0)$$

$$= \int_0^T P(\mathbf{X}_1 = t \mid s) \int_t^T P(\mathbf{X}_2 = x \mid n) dx\, dt + b P(\mathbf{X}_1 = T \mid s) P(\mathbf{X}_2 = T \mid n)$$

$$= \int_0^T \mu(e^{-\mu t} - e^{-\mu T}) [(e^{-\nu t} + \nu t e^{-\nu T} - (1 + \nu T)e^{-\nu T}] dt + b(1 + \mu T)$$
$$\times (1 + \nu T) e^{-(\mu+\nu)T}$$

$$= \frac{\mu}{\mu + \nu} + \left(\frac{\nu}{\mu} - 1 - \nu T\right) e^{-\nu T} - \frac{\mu}{\nu} e^{-\mu T} + e^{-(\mu+\nu)T}\left[1 + \frac{\mu}{\nu} - \frac{\nu}{\mu}\right.$$
$$\left. - \frac{\mu}{\mu + \nu} + \mu T + \frac{\mu\nu T^2}{2} + b(1 + \mu T)(1 + \nu T)\right]. \quad [41]$$

Observe that as $T \rightarrow \infty$, the model reduces to the so-called unbiased choice model suggested by various authors (for some history see Luce, 1959, 1963a; see

the section on "Discrimination of Intensity" for a generalization).

If we assume that varying b generates the forced-choice ROC curve, then it is

given by

$$p(1 \mid s,n) - p(1 \mid n,s)$$

$$= \frac{\mu - \nu}{\mu + \nu} + \left(2\frac{\mu}{\nu} - 1 - \nu T\right)e^{-\nu T}$$

$$- \left(2\frac{\nu}{\mu} - 1 - \mu T\right)e^{-\mu T}$$

$$+ e^{-(\mu+\nu)T}\left[2\frac{\mu}{\nu} - 2\frac{\nu}{\mu}\right.$$

$$\left. - \frac{\mu - \nu}{\mu + \nu} + (\mu - \nu)T\right]. \quad (42)$$

Note that this is a straight line with slope 1 as has been obtained by similar considerations in the threshold theories (Atkinson, 1963; Krantz, 1969; Luce, 1963a, 1963b; Norman, 1964). Here, in effect, the threshold is not in the subject but in the limited listening interval of the experimental design.

It is important to realize that according to the present theory, the method of generating the ROC curve differs in the yes–no and forced-choice experiments. In the former, a response criterion is varied much as in TSD; in the latter, a response bias is applied to inherently ambiguous information much as in threshold theories. In both cases, the limited period of signal presentation leads to an effective threshold which is reflected in the ROC curves. Eliminating this threshold by having very long intervals $(T \to \infty)$ drives the yes–no ROC curve toward a power function and the forced-choice one toward a single point. The present model suggests that different aspects of TSD, threshold, and choice theories will be evident depending on details of the experimental situation.

SIMPLE REACTION TIMES WITH RANDOM FOREPERIODS

Consider next a simple RT experiment in which there is a signal on every trial and the subject is asked to respond as soon as he detects it. Unlike the previous designs, there is only one kind of trial and only one response. In order to make the task difficult and to insure that the subject
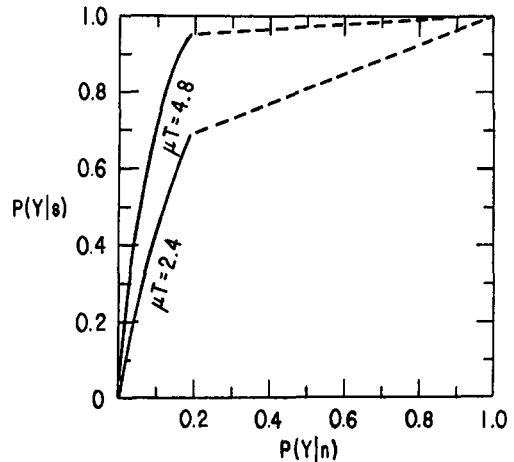

FIG. 6. ROC curve for decision rule based on first IAT within signal interval. (T is the duration of the signal interval, μ the rate parameter associated with the signal. The noise alone rate parameter, ν, times the signal interval was .8 for these calculations, $\nu T = .8$.)

really waits for the signal, it is common experimental practice to make the time from the beginning of the trial (warning signal) to the onset of the reaction signal a random variable. (In our experiments, we have used a discrete approximation to an exponential density because in that case the length of the wait provides absolutely no information about when the signal will occur; some experiments have used other, less desirable distributions.) The only unconventional features of the design we shall study is that the signals are response terminated. This seems appropriate when weak signals are used; it is usually not done with the intense signals used in most RT experiments.

The analysis below was carried out because we had run such an experiment to test a previous theory. It is included here, even though in important aspects it is incomplete, because the experiment is inherently interesting and because, in contrast to the above models, the analysis is mathematically more difficult. The only case that we work out at all fully is $J = \kappa = 1$. To analyze data, however, we really need the $J > 1$ case. The only model we have worked out, which is somewhat different from those above, assumes that

the $J = 1$ model applies to each of the channels individually, and that the outputs of these channels all feed into a single channel to which the $J = 1$ model again applies. No work has been done on $\kappa > 1$.

Let $\mathbf{S}$ denote the random time of signal onset after the beginning of the trial, that is, the foreperiod, and let $\mathbf{R}$ denote the time of the response, again measured from the beginning of the trial. Our problem is to derive the distribution of $\mathbf{R}$ under the decision rule that a response is initiated at the end of the first IAT that does not exceed β. The difficulty of the problem lies in the fact that we work not with the first IAT, as in the yes–no designs, but rather with the first IAT that meets a criterion. This difficulty appears to be inherent in any situation where the possible onset of the signal is not specified experimentally.

Analytically, the distribution of $\mathbf{R}$ is composed of its separate distributions at fixed foreperiods, each component distribution being weighted by the probability of that foreperiod. So we first examine the case of a fixed foreperiod, that is, $\mathbf{S} = x$. We, therefore, consider a Poisson process with parameter ν during the interval $[0,x)$ and with parameter μ ($>\nu$) during the interval $[x,\infty)$. Let $\mathbf{Y}_i$ denote the time of the ith event in this process, where $\mathbf{Y}_{i+1} \geq \mathbf{Y}_i$. Define the decision random variable $\mathbf{D}$ as follows:

$\mathbf{D} = t$ if and only if for some integer j,

$$\mathbf{Y}_j = t, \quad \mathbf{Y}_j - \mathbf{Y}_{j-1} \leq \beta,$$

and for all $i = 1, \cdots, j - 2$,

$$\mathbf{Y}_{i+1} - \mathbf{Y}_i > \beta.$$

That is to say, $\mathbf{D}$ is the wait until the end of the first IAT that is $\leq \beta$. We introduce the notation

$$l(x,t) = P(\mathbf{D} = t \mid \mathbf{S} = x) \qquad [43]$$

for the distribution of decision latencies. It is also convenient theoretically to introduce

$$l_0(x,t)$$
$$= P(\mathbf{D} = t \mid \mathbf{S} = x \text{ and } \mathbf{Y}_1 = 0), \quad [44]$$

which is the same density subject to the condition that the first event occurs at time 0.

Although most RT data are given in the form of approximations to the following conditional density

$$f_{R-S}(t) = P(\mathbf{R} - \mathbf{S} = t \mid \mathbf{R} \geq \mathbf{S})$$
$$= \int_0^\infty P(\mathbf{R} - \mathbf{S} = t \mid \mathbf{S} = x, \mathbf{R} \geq x)$$
$$\times P(\mathbf{S} = x)dx,$$

it is not sufficient simply to study $l(x,t)$ for $t \geq x$. The reason is that the observed response time density is the convolution of the decision density with the residual one, and so some responses following signal onset arise from decisions reached before it.

Our problem is to derive an equation for $l(x,t)$. This is a rather more difficult task than its simple statement suggests. The only relevant literature which we have found is Gilbert and Pollak (1957) who studied the case $x = 0$; they called it a Poisson coincidence counter. To some degree we will repeat a portion of their work, but it is more difficult than it is worth not to make our treatment of our particular problem self-contained.

Expressions for $l_0(0,t)$

The key to understanding the process is, first, to understand the simplest case in which there is only one Poisson process and the first pulse is at the origin, that is, to solve for l_0. In Appendix C we show that l_0 satisfies

$$l_0(0,t) = \begin{cases} \mu e^{-\mu t}, & 0 \leq t \leq \beta \\ \mu e^{-\mu t} \int_0^{t-\beta} e^{\mu \nu} l_0(0,y)dy, & t > \beta, \end{cases} \qquad [45]$$

which is an integral equation of retarded type with a lag of β. We can solve it in three ways (see Appendix C), each of which may prove useful at one time or another.

First, we have an explicit series expansion; we have found this expansion least useful.

$$l_0(0,t) = \mu e^{-\mu t} \sum_{i=0}^{m-1} a(m-1-i)$$

$$\times \frac{\mu^i [t - (m-1)\beta]^i}{i!},$$

$$(m-1)\beta \le t < m\beta \quad [46]$$

where

$$a(0) = 1, \quad a(1) = 0,$$

$$a(m) = \sum_{i=1}^{m-1} \sum_{k=1}^{i} a(i-k) \frac{(\mu\beta)^k}{k!},$$

$$\text{for } m > 1. \quad [47]$$

Second, we work out an approximate solution for small β. The trick is to make a Taylor's series expansion of the integral in Equation 45 about t in powers of β and retain only the linear term. The free constant is determined by the fact that s_0 is a density. This yields the approximation

$$l_0(0,t) \cong \begin{cases} \mu e^{-\mu t}, & 0 \le t \le \beta \\ \mu' e^{-\mu' / \mu} e^{-\mu' t} & t > \beta \end{cases} \quad [48]$$

where

$$\mu' = \frac{\mu^2 \beta}{1 + \mu\beta}. \quad [49]$$

And third, we can find the Fourier transform of $l_0(0,t)$; it is surprisingly simple:[10]

$$L_0(0,\omega) = \int_0^\infty e^{-i\omega t} l_0(0,t) dt$$

$$= \frac{\mu[1 - e^{-(i\omega+\mu)\beta}]}{i\omega + \mu - \mu e^{-(i\omega+\mu)\beta}}. \quad [50]$$

---

[10] As J. W. Tukey suggested to us, a simple direct argument leads to Equation 50. The process consists of $n = 1, 2, \cdots$, independent drawings from the exponential distribution, the first $n-1$ of which are $> \beta$ and the nth is $\le \beta$. As is easily computed, these two truncated densities have transforms

$$U(\omega) = \frac{\mu e^{-i\omega\beta}}{\mu + i\omega} \quad \text{and} \quad V(\omega) = \frac{\mu[1 - e^{-(\mu+i\omega)\beta}]}{(1 - e^{-\mu\beta})(\mu + i\omega)}$$

and probabilities of occuring $e^{-\mu\beta}$ and $1 - e^{-\mu\beta}$, respectively. Thus,

$$L_0(0,\omega) = \sum_{n=1}^{\infty} [e^{-\mu\beta} U(\omega)]^{n-1} (1 - e^{-\mu\beta}) V(\omega)$$

$$= \frac{(1 - e^{-\mu\beta}) V(\omega)}{1 - e^{-\mu\beta} U(\omega)},$$

which is Equation 50.

Results for $l(x,t)$ *when* $x \le \beta$

We now use these expressions for l_0. First, we find the Fourier transform of l (see Appendix C):

$$L(x,\omega) = \int_0^\infty e^{-i\omega t} l(x,t) dt$$

$$= \left(\frac{\nu}{i\omega + \nu}\right)^2 \{1 - [(i\omega + \nu)x + 1]$$

$$\times e^{-(i\omega+\nu)x}\} + \frac{e^{(\mu-\nu)x}}{i\omega + \mu} \Big\{ L_0(0,\omega)$$

$$\times [\nu + (\mu - \nu)e^{-(i\omega+\mu)x}]$$

$$+ \mu\nu x e^{-(i\omega+\mu)x} - \frac{\mu\nu}{i\omega + \mu}$$

$$\times [1 - e^{-(i\omega+\mu)x}] \Big\}, \quad [51]$$

where $L_0(0,\omega)$ is given in Equation 50.

Assuming for the moment that we have estimates of the parameters ν, μ, and β, we can use Equation 51 as follows. With a fixed foreperiod, the response density is a convolution of the decision time with the residual time, so its Fourier transform $F(x,\omega)$ can be written

$$F(x,\omega) = L(x,\omega) R(\omega).$$

Or, to use more data, if we have a distribution g of foreperiods,

$$F(\omega) = \int_0^\beta g(x) F(x,\omega) dx$$

$$= \int_0^\beta g(x) L(x,\omega) dx R(\omega).$$

This permits us to calculate R and so r from the data using Equation 51. The only difficulty is in estimating the parameters, with which we deal below.

For $x \le \beta$, we can compute the MRT from a signal onset at time x by using Equation 51 and the fact that

$$\int_0^\infty t f(t) dt = i \lim_{\omega \to 0} \frac{dF(\omega)}{d\omega}.$$

Carrying out that computation,

$$\text{MRT}_x = \bar{r} + i \lim_{\omega \to 0} \frac{dL(x,\omega)}{d\omega} - x$$

$$= \bar{r} + \frac{2}{\nu}(1 - e^{-\nu x}) + (A - Bx)e^{-\nu x}$$

$$+ Ce^{(\mu - \nu)x} - x, \quad x \leq \beta, \quad [52]$$

where

$$A = \left(\frac{\mu - \nu}{\mu^2}\right)\left(\frac{1}{1 - e^{-\mu\beta}}\right) + \frac{1}{\mu} + \frac{\nu}{\mu^2}$$

$$B = 1 - \frac{\nu}{\mu}$$

$$C = \frac{\nu}{\mu^2}\left(\frac{1}{1 - e^{-\mu\beta}} - 1\right).$$

For $\nu = 0$ or $x = 0$,

$$\text{MRT}_x = \bar{r} + \frac{1}{\mu}\left(\frac{1}{1 - e^{-\mu\beta}} + 1\right) \quad [53]$$

and

$$\left(\frac{\partial \text{MRT}_x}{\partial x}\right)_{x=0} = -\frac{\nu}{\mu} < 0.$$

It is evident that as the mean foreperiod is increased, the subject can maintain a fixed false alarm rate only by decreasing β. Since this increases both A and C, MRT_x, $x \leq \beta$, must increase with increasing mean foreperiod. This was observed in Green and Luce (1971). Unfortunately, we cannot check their observations in greater detail until we know more about $l(x,t)$ for $x > \beta$.

MRT Data

Assuming that the noise is negligible, so Equation 53 applies, and recalling from Equation 17 that the mean (or geometric mean) MEs are proportional to μ, that is, $\text{ME} = D\mu$, then

$$\text{MRT} \cong \bar{r} + \begin{cases} \dfrac{2D}{\text{ME}}, & \text{for } \mu \text{ large} \\[2mm] \dfrac{D}{\text{ME}} + \dfrac{D^2}{\text{ME}^2\beta}, & \text{for } \mu \text{ small} \\ \qquad (1 - e^{-\mu\beta} \cong \mu\beta). \end{cases}$$

Vaughan, Costa, and Gilden (1966) showed that the data of several authors for visual signals of fixed duration are consistent with the formula

$$\text{MRT} = \bar{r} + D'/I^{\frac{1}{3}},$$

where I is luminance. As they point out, it is noteworthy that "subjective brightness," as determined by the method of ME, grows as $I^{\frac{1}{3}}$ (J. C. Stevens & Stevens, 1963; S. S. Stevens, 1957; S. S. Stevens & Galanter, 1957). Assuming that the magnitude estimates are proportional to $I^{\frac{1}{3}}$, then at least for moderately large μ we would expect Vaughan et al.'s result. The only complication in this argument is that they obtained the data using exceedingly brief signals—in their own data, $T = 10$ μsec. (in those of others $T = 55$ msec. to 575 msec.). For these data to be compatible with our theory, it is necessary to suppose that the effect of a very brief visual signal on the neural pulse train is much longer than the signal's duration.

For loudness, McGill (1960) has interpreted his data as showing MRT is linear in $\text{ME}^{-\frac{1}{3}}$. Taking the data reported in McGill (1963), unpublished data of J. G. Snodgrass (personal communication, 1962), and those of Chocholle (1940) and choosing $\bar{r}$ to be 5 msec. smaller than the MRT to the loudest signal, we obtain the curves shown in Figure 7. Since Chocholle did not obtain MEs, we have plotted his data as if the magnitude exponent were .3 (his other two subjects give nearly identical results). These plots seem to support the theory; however, whether or not one says the slopes are -1, as predicted, depends on exactly what choice one makes for $\bar{r}$. Altering it by only a few milliseconds has an enormous effect on the apparent slope. McGill made a different choice from ours. Chocholle's data, which cover the range down to the subject's threshold, exhibit the predicted steepening of slope at low intensities.

Estimation of Parameters

We need schemes to estimate ν, μ, and β. As we shall see, we have satisfactory methods for estimating ν and β, but not for μ except when β is small. For other β we do not presently have a method for estimating μ.
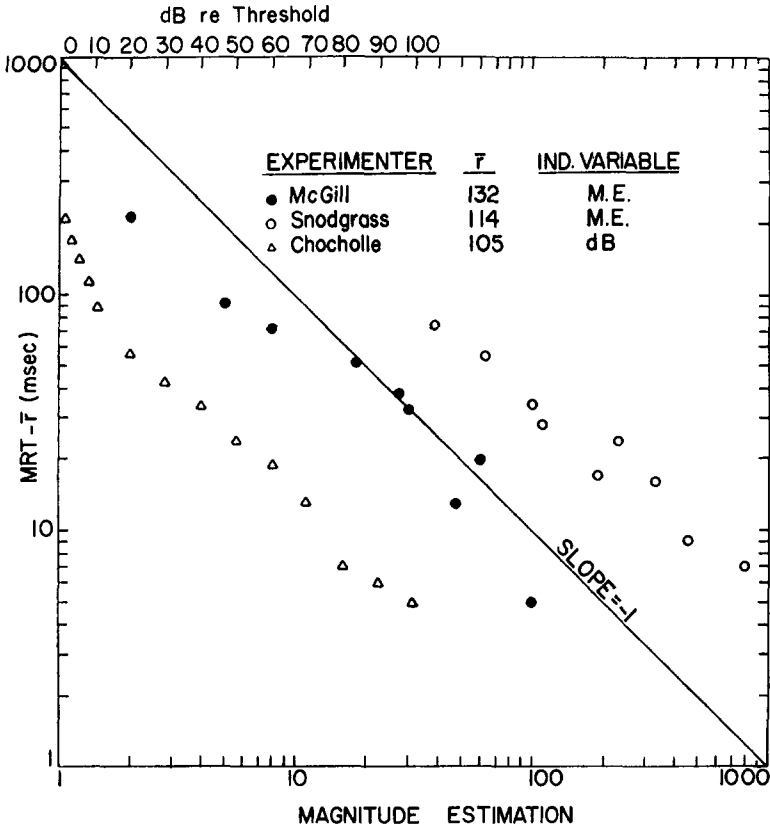
FIG. 7. Plots of MRT $-\bar{r}$ versus ME of loudness of 1000-Hz. tones from three experiments. (For Chocholle's data, it is assumed that ME $= \alpha I^{.3}$. In each case, F $=$ MRT$_{min} - 5$ [in milliseconds]; for the Snodgrass data, the loudest signal [ME $= 1500$] is not on the graph.)

On the assumption of a known small β, then to estimate μ it is sufficient to estimate μ' and use Equation 49. To estimate μ' we use the approximate expression for $l_0(0,t)$, Equation 48 (which is where the assumption of small β enters), in equations for $l(x,t)$, convolute the result with r, and integrate that over an arbitrary distribution of foreperiods. As we show in Appendix C, the resulting density of reaction times has tails of the following form:

$$f_{R-S}(t) = P(R - S = t | R \geqslant S)$$
$$\cong Ae^{-\mu t} + Be^{-\mu't},$$
$$t > \tau + 2\beta, \quad [54]$$

where μ' is given by Equation 49 and A and B depend on ν, μ, r, and the distribu-

tion of foreperiods. Observe that

$$\frac{\mu'}{\mu} < \frac{\mu\beta}{1 + \mu\beta},$$

and so for β sufficiently small, μ' is appreciably smaller than μ. Thus, $e^{-\mu't}$ is the dominant term in Equation 54 and so the tail of the distribution of RTs allows us to estimate μ'.

To get at the parameter ν, consider $l(x,t)$ for $t \leq x$. It can be shown (Appendix C) that $l(x,t) = l(t,t)$ for $t \leq x$ and that it satisfies the following integral equation of retarded type:

$$l(t,t) = \begin{cases} \nu^2 t e^{-\nu t}, & t \leq \beta \\ \nu e^{-\nu t}\left(\int_0^{t-\beta} e^{\nu z} l(z,z)dz + \nu\beta\right), \\ & t > \beta. \quad [55] \end{cases}$$

To calculate the Fourier transform, we note that $l(t,t) = l(\infty,t)$, and the latter is simply the convolution of $\nu e^{-\nu t}$ with $l_0(\infty,t)$. Therefore,

$$L(\omega) = L(\infty,\omega) = \frac{\nu}{\nu + i\omega} L_0(\infty,\omega).$$

However, $L_0(\infty,\omega)$ is just $L_0(0,\omega)$ with μ replaced by ν, so by Equation 50,

$$L(\omega) = \frac{\left(\dfrac{\nu}{i\omega + \nu}\right)^2 [1 - e^{-(i\omega+\nu)\beta}]}{1 - \dfrac{\nu}{i\omega + \nu} e^{-(i\omega+\nu)\beta}}. \quad [56]$$

Moreover, if we assume that for $t > \tau'$

$$l(t,t) \cong Ce^{-\nu' t}, \quad [57]$$

then direct substitution into Equation 55, collecting terms in $e^{-\nu t}$ and $e^{-\nu' t}$, and choosing ν' and C so that the equation is satisfied yields

$$\nu' = \nu[1 - e^{-(\nu-\nu')\beta}], \quad C = \nu\beta(\nu - \nu'). \, [58]$$

Let f_R denote the density of response times for which the response precedes the signal onset. If we assume that the density of foreperiods is $\lambda e^{-\lambda t}$, then

$$f_R(t) = Ke^{-(\lambda+\nu')t}, \quad t \geq \tau' + \tau, \quad [59]$$

where K depends on ν, β, and r. Thus, we can estimate ν' from the tail of f_R.

In order to estimate ν and β, we need another expression relating them. We turn to that next.

False Alarm Probability

If we use an exponential distribution of foreperiods, the probability of a false alarm is easily seen to be given by

$$P(R < S)$$
$$= \int_0^\infty e^{-\lambda x} \int_0^x l(y,y) r(x - y) dy dx,$$

which is simply the LaPlace transform of the convolution of l and r, and so it is the product of their respective LaPlace transforms. By the same technique as used to get the Fourier transform of $l(t,t)$ in

Equation 56, we obtain

$$P(R < S)$$
$$= \frac{\left(\dfrac{\nu}{\lambda + \nu}\right)^2 [1 - e^{-(\lambda+\nu)\beta}]}{1 - \dfrac{\nu}{\lambda + \nu} e^{-(\lambda+\nu)\beta}} R_{-\lambda}, \quad [60]$$

where

$$R_{-\lambda} = \int_0^\infty e^{-\lambda t} r(t) dt. \quad [61]$$

Observe that by solving Equation 58 for β and substituting it in Equation 60, we obtain

$$\frac{P(R < S)}{R_{-\lambda}} = \frac{1 - f(x,\alpha)}{(1 + \alpha x)[1 + \alpha x - f(x,\alpha)]},$$

where

$$x = \nu'/\nu$$
$$\alpha = \lambda/\nu'$$
$$f(x,\alpha) = (1 - x)^{(1+\alpha x)/(1-x)}.$$

This function is plotted in Figure 8. It can be used as follows: Since λ is determined by the experimenter and ν' can be estimated by Equation 58, α is known. Observe that for λ sufficiently small,

$$R_{-\lambda} \cong 1 - \lambda \bar{r}.$$

We can determine $\bar{r}$ from a very intense signal, so $P(R < S)/R_{-\lambda}$ is also known. Thus, x is determined, from which we obtain ν, and so β is determined from the plot of $\beta\nu$ (Equation 58) also included in Figure 8.

Moments of the Response Time Distribution

If the signal onset is exponentially distributed, the overall response time density is

$$\int_0^\infty \lambda e^{-\lambda x} l(x,t) dx.$$

Its Fourier transform is

$$L_\lambda(\omega) = \int_0^\infty e^{-i\omega t} \int_0^\infty \lambda e^{-\lambda x} l(x,t) dx dt$$
$$= \int_0^\infty \lambda e^{-\lambda x} L(x,\omega) dx,$$

where L is given by Equation 51 for $x \leq \beta$. If an explicit expression could be obtained
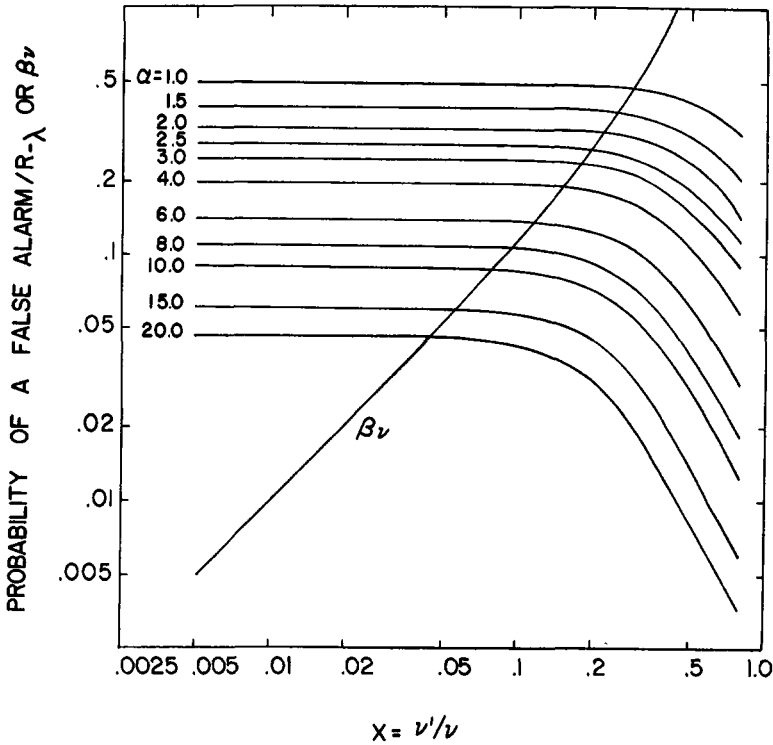
FIG. 8. Plots of $P(\mathbf{R} < \mathbf{S})/R_{-\lambda}$ versus $x = \nu'/\nu$ with $\alpha = \lambda/\nu'$ as a parameter and of $\beta\nu$ versus x. (Assuming $P(\mathbf{R} \leq \mathbf{S})/R_{-\lambda}$ and α known, this plot may be used to estimate ν and β.)

for $x > \beta$, then the moments could be obtained by taking derivatives of $L_\lambda(\omega)$ and letting $\omega \to 0$. For example, the mean would be given by

$$\lim_{\omega \to 0} i \frac{dL_\lambda(\omega)}{d\omega}.$$

In principle, this gives us all the additional equations needed to estimate the parameters without invoking any approximations.

Estimates of μ for Sound Intensity

It is easy to estimate μ_{max}/μ_{min} from the empirical fact that, roughly, I_{max}/I_{min} is 10^{10}. So, from the ME exponent of .3, we conclude μ_{max}/μ_{min} is about 10^3. It should be realized that this estimate is considerably larger than the observed range of pulse rates in the peripheral fibers (Grossman & Vierstein, 1961; Kiang, 1965; Rodieck et al., 1962). A range of 10^2 is unusual; 10 is much more common. This suggests either some sort of transfer from one

neural path to another as intensity increases or that the central system has considerably more dynamic range.

We do not know of psychophysical data that permit us to estimate either μ_{min} or μ_{max}. The nearest that we can come to it, which illustrates the method we would use to estimate μ_{min} if the data were available, is to estimate the value ν, corresponding to the noise background used in our experiments, namely, a noise power density of 40 db. We use two methods. The first probably underestimates ν because it assumes $J = 1$. It is based on the analysis of the random foreperiod RT experiment. We estimate ν and β by the method suggested following Equation 60. According to Green and Luce (1971), in an experiment with mean exponential foreperiod of 2 seconds ($\lambda = .5$) and with 10,000 observations collected from each of two subjects, the tail of f_R (Equation 59) is exponential, $\nu' \cong .07$, and $P(\mathbf{R} < \mathbf{S}) \cong .1$.

If we assume, as is plausible, $\hat{r} = .2$, then $R_{-\lambda} = 1 - (.5)(.2) = .9$. Solving in Figure 8, $\nu \cong .35$ and $\beta \cong .8$.

The second method is based on fixed-interval yes–no data and on the assumption that the response rule ignores the termination of the interval. In that case, the noise equation is the same as in the response terminated case, Equation 26. With a reasonably strict criterion, it is probably safe to assume $\kappa = 1$, $\beta \leq .2$, and that $p(Y|n) \cong .1$. If so, using the normal approximation to the gamma, we see by Equation 30

$$J^{\frac{1}{2}}(\nu\beta - 1) = z(n) \cong -1.30.$$

Using our estimate of $J \cong 27$, we obtain $\nu \geq 3.72$. Note that this corresponds well with the estimate from Grossberg's (1968) MRT data for brightness.

With the latter estimate, a signal 65 db. above the noise would have a value of μ/ν of about $10^{6.5(.3)} = 90$, so $\mu \geq 335$, For $T = \frac{1}{2}$ second, $\mu T \geq 167 > 65 \cong K$, hence these numbers are consistent with our definition of a strong signal.

According to the data of Rose et al. (1967), the representation of signals of frequencies not greater than 2500 Hz. in the peripheral neurons is quantized in the period of the signal. Thus, a maximum rate for a 1000-Hz. signal would be a pulse every millisecond, that is, a value of 1000 for μ. This does not appear grossly inconsistent with the psychophysical estimates.

The order of magnitude difference in our estimates of ν, .35 and 3.72, probably results from the fact that we have incorrectly assumed $J = 1$ in analyzing the random foreperiod experiment. To show that this may account for the discrepancy, consider the following $J > 1$ model for this experiment. Each of the J parallel channels processes pulses as we have assumed; however, whenever two successive pulses are less than δ apart, it triggers not a response, but rather a pulse on a common channel. If δ is small, say a few milliseconds, this could be done by energy summation at a synapse with a threshold. The common channel, then, carries the superposition of the outputs of the J parallel channels. When there is no signal, the pulses on the common channel should not cluster; when there is, they should cluster. Moreover, if J is sufficiently large, the process on the common channel is approximately Poisson, even though the recurrent events on the parallel channels are not (see Cox, 1962, pp. 71–79). Finally, we assume that the common channel triggers a response according to the $J = 1$ model with a criterion $\beta \neq \delta$.

From Equation 50, we obtain

$$1/\nu(1 - e^{-\nu\delta})$$

as the mean time between successive outputs of each of the parallel channels. Thus, the intensity parameter of the common channel is

$$\nu' = J\nu(1 - e^{-\nu\delta}) \cong J\nu^2\delta.$$

So, if $\nu = 3.72$, $J = 27$, and $\delta = .001$, then $\nu' = .373$, which seems to account for the order of magnitude difference in our estimate.

According to the arguments leading to Equations 54 and 59, the tails of both the false alarm and RT densities should be exponential in a simple RT experiment with an exponential distribution of foreperiods. Luce and Green (1970) verified this prediction, and they estimated[11] the parameters μ' and ν' for various signal intensities. Let μ_1 and ν_1 denote the corresponding true single-channel values which are related to μ' and ν' by Equations 49 and 58, respectively. Finally, let μ and ν be the underlying parameters on the several parallel channels, as just derived. If $\mu_1\beta$ is small

[11] The cut point used in analyzing these data was 500 msec. Recall that Equation 54 was derived for $t > \tau + 2\beta$. We are probably safe in assuming τ is not greater than 300 msec.; this is based on simple RT distributions to intense signals, which according to this theory is dominated by τ. So the above inequality would hold if β were less than 100 msec. However, using the method for estimating β outlined in connection with Figure 8, we estimate β for these data to be about 1 second. Therefore, we have no reason to believe our approximations are applicable. In spite of that, we continue the calculation as if they are justified because the results seem consistent with earlier ones.
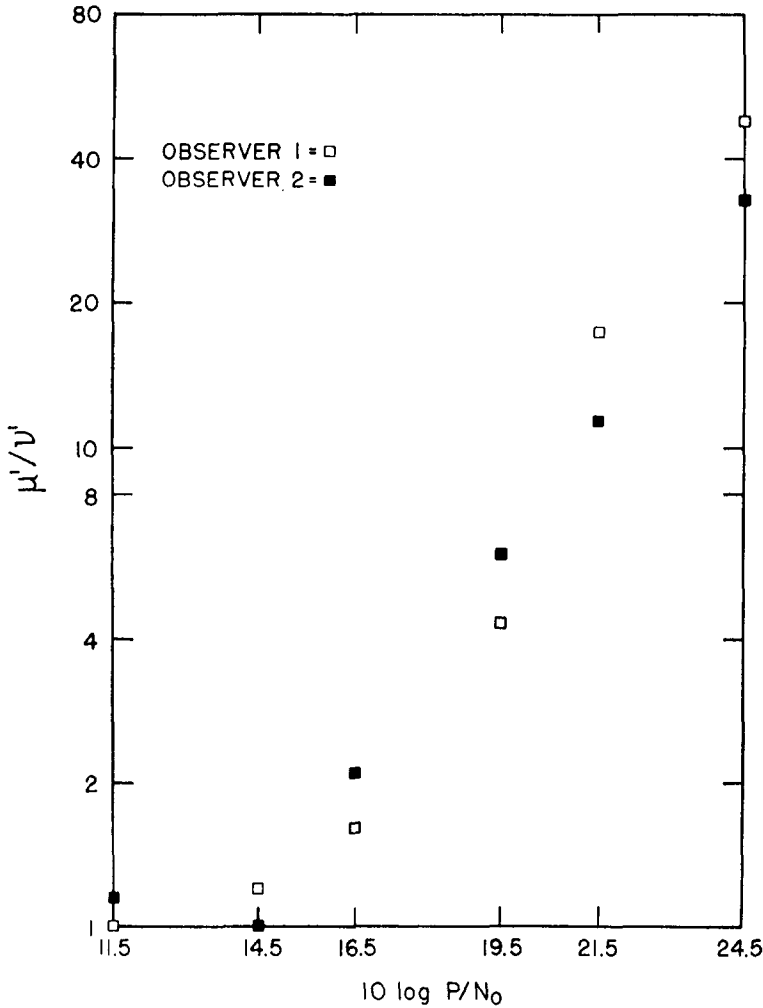
FIG. 9. Estimated μ'/ν' versus the ratio of signal power to noise density in decibels. (The two Poisson parameters were estimated from the tails [> 500 msec.] of RT densities. This is Figure 9 of Luce & Green, 1970.) (Reprinted from an article by R. D. Luce and D. M. Green in *Perception and Psychophysics*, Vol. 7. Copyrighted by Psychonomic Journals, Inc., 1970.)

relative to 1, then

$$\frac{\mu'}{\nu'} \cong \frac{\mu_2{}^2\beta}{1+\mu_1\beta} \cdot \frac{1+\nu_1\beta}{\nu_2{}^2\beta} \cong \left(\frac{\mu_1}{\nu_1}\right)^2 \cong \left(\frac{\mu}{\nu}\right)^4.$$

Figure 9 presents the empirical plot of μ'/ν' versus signal-to-noise in decibels for two subjects. We see that to a good approximation μ'/ν', and so by what we just showed μ/ν, is a power function of $I(s)/I(n)$, confirming again one of our basic postulates. Moreover, taking into account the factor of 4, the estimated exponent γ

is .4, which is well within the range of values found by magnitude estimation of the intensity of pure tones.

For a particular signal-to-noise ratio, we can also estimate μ/ν from response terminated yes–no ROC data. For example, we have found that $10 \log P/N_0 = 16.5$ db. yields a value of about 1 for d', that is, $z(s) = 1$ and $z(n) = 0$. From Equation 32, we see that

$$\frac{\mu}{\nu} = 1 + 1/J^{\frac12},$$

which with our estimate of $J \cong 27$ yields $\mu/\nu = 1.193$. By the above argument, the observed value of μ'/ν' should be $(1.193)^4 = 1.85$, which falls between the two data points. If we reverse the argument, this provides an independent estimate of J which agrees with the earlier one from ME data.

SUMMARY

The postulates of the theory are: signal intensity is transformed into J identical Poisson processes on parallel channels; the (intensity) parameter of these Poisson processes is a power function, with exponent γ, of signal intensity; decisions are based on the IATs of pulses, which are estimates of the Poisson parameter; up to K IATs per channel can be collected and stored prior to reaching a decision; and observed responses are delayed beyond the decision time by the addition of a residual latency which is no longer than τ seconds.

These ideas were worked out in some detail for various standard types of experiments, and a number of well-known results were shown to follow, several postulates were tested, and estimates were made of the several parameters.

Aside from neurophysiological data on peripheral nerves, the strongest evidence for the Poisson assumption comes from the exponential decay of RT densities. Support for the power function came from three direct sources—ME data, the nature of the decrease in MRT with signal intensity, and a plot of Poisson parameters (estimated from the tails of RT densities) versus signal intensity—and one indirect one—the prediction of Weber's law. It is gratifying that all three estimates of the exponent γ were about the same, .3–.4 for sound intensity. Actual values of the Poisson parameters, based on yes–no and RT data, suggest a value of ν, corresponding to the noise background, of about 3–4, or a mean IAT of 250–350 msec. A signal-to-noise ratio corresponding to a d' of 1 yields a ratio of μ/ν of about 1.2.

The estimates of the constants J and K are less secure, and future work should attempt to find additional ways to estimate them. From Weber's fraction, we concluded that for 1000-Hz. pure tones, JK has a value of about 1750. Two arguments, one based on the ratio of the mean to standard deviation of MEs and another on a yes–no ROC curve and a previous estimate of μ/ν, suggested that J is about 27, in which case K is about 65. Finally, we estimated τ to be about 300 msec. on the ground that simple RTs to intense signals are very nearly the residual latency.

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.

BARLOW, H. B. Increment thresholds at low intensities considered as signal/noise discriminations. *Journal of Physiology*, 1957, 136, 469–488.

BATHER, J. A. On a quickest detection problem. *Annals of Mathematical Statistics*, 1967, 38, 711–724.

BELLMAN, R., & COOKE, K. L. *Differential-difference equations.* New York: Academic Press, 1963.

BOUMAN, M. A. History and present status of quantum theory in vision. In W. A. Rosenblith (Ed.), *Sensory communication.* Cambridge: M.I.T. Press & Wiley, 1961.

BRAIDA, L. D., & DURLACH, N. I. Intensity perception, II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, 1972, 52, in press.

BRUGGE, J. F., ANDERSON, D. J., HIND, J. E., & ROSE, J. E. Time structure of discharges in single auditory nerve fibers of the squirrel monkey in response to complex periodic sounds. *Journal of Neurophysiology*, 1969, 32, 386–401.

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

CHOCHOLLE, R. Variations des temps de réaction auditifs en fonction de l'intensité à diverses fréquences. *L'Année Psychologique*, 1940, 41, 5–124.

CHRISTIE, L. S., & LUCE, R. D. Decision structure and time relations in simple choice behavior. *Bulletin of Mathematical Biophysics*, 1956, 18, 89–112.

COX, D. R. *Renewal theory.* London: Metheun, 1962.

CREELMAN, C. D. Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, 1961, 34, 582–593. Reprinted in J. A. Swets (Ed.), *Signal detection and recognition by human observers.* New York: Wiley, 1964.

DURLACH, N. I., & BRAIDA, L. D. Intensity perception, I. Preliminary theory of intensity resolu-

tion. *Journal of the Acoustical Society of America*, 1969, **46**, 372–383.

EDWARDS, W. Optimal strategies for seeking information: Models for statistics, choice reaction times and human information processing. *Journal of Mathematical Psychology*, 1965, **2**, 312–329.

EGAN, J. P., GREENBERG, G. Z., & SCHULMAN, A. I. Operating characteristics, signal detectability, and the method of free response. *Journal of the Acoustical Society of America*, 1961, **33**, 993–1007.

EL'SGOL'TS, L. E. *Introduction to the theory of differential equations with deviating arguments.* San Francisco: Holden-Day, 1966.

ESTES, W. K. A random walk model for choice behavior. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford: Stanford University Press, 1960.

GILBERT, E. N., & POLLAK, H. O. Coincidences in Poisson patterns. *Bell System Technical Journal*, 1957, **36**, 1005–1033.

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

GREEN, D. M., & LUCE, R. D. Detection of auditory signals presented at random times, III. *Perception and Psychophysics*, 1971, **9**, 257–268.

GREEN, D. M., & LUCE, R. D. Speed-accuracy trade-off in auditory detection. In S. Kornblum (Ed.), *Attention and performance.* Vol. 4. New York: Academic Press, 1972, in press.

GREEN, D. M., & MCGILL, W. J. On the equivalence of detection probabilities and well-known statistical quantities. *Psychological Review*, 1970, **77**, 294–301.

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

GRICE, G. R. Stimulus intensity and response evocation. *Psychological Review*, 1968, **75**, 359–373.

GROSSBERG, M. The latency of response in relation to Bloch's law at threshold. *Perception and Psychophysics*, 1968, **4**, 229–232.

GROSSMAN, R. G., & VIERSTEIN, L. J. Discharge patterns of neurons in cochlear nucleus. *Science*, 1961, **134**, 99.

HIND, J. E., ANDERSON D. J., BRUGGE, J. F., & ROSE, J. E. Coding of information pertaining to paired low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 1967, **30**, 794–816.

JOHN, I. D. A statistical decision theory of simple reaction time. *Australian Journal of Psychology*, 1967, **19**, 27–34.

KENDALL, M. G., & STUART, A. *The advanced theory of statistics: I.* New York: Hafner, 1958.

KIANG, N. Y-S. *Discharge patterns of single fibers in the cat's auditory nerve.* Cambridge: M.I.T. Press, 1965.

KIANG, N. Y-S. A survey of recent developments in the study of auditory physiology. *Annals of Otology, Rhinology and Laryngology*, 1968, **77**, 656–676.

KINCHLA, R. A., & SMYZER, F. A diffusion model of

perceptual memory. *Perception and Psychophysics*, 1967, **2**, 219–229.

KOHFELD, D. L. Effects of the intensity of auditory and visual ready signals on simple reaction time. *Journal of Experimental Psychology*, 1969, **82**, 88–95.

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

KRISTOFFERSON, A. B. Successiveness discrimination as a two-state quantal process. *Science*, 1967, **158**, 1337–1339. (a)

KRISTOFFERSON, A. B. Attention and psychophysical time. *Acta Psychologica*, 1967, **27**, 93–100. (b)

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

LAMING, D. R. J. *Information theory of choice-reaction times.* New York: Academic Press, 1968.

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

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. (a)

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

LUCE, R. D. A model for detection in temporally unstructured experiments with a Poisson distribution of signal presentations. *Journal of Mathematical Psychology*, 1966, **3**, 48–64.

LUCE, R. D., & GREEN, D. M. Detection of auditory signals presented at random times, II. *Perception and Psychophysics*, 1970, **7**, 1–14.

LUCE, R. D., & MO, S. S. Magnitude estimation of heaviness and loudness by individual subjects: A test of a probabilistic response theory. *British Journal of Mathematical and Statistical Psychology*, 1965, **18**, 159–174.

MCGILL, W. J. Loudness and reaction time: A guided tour of the listener's private world. *Proceedings of the XVI International Congress of Psychology, 1960.* Amsterdam: North-Holland, 1960.

MCGILL, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 1. New York: Wiley, 1963.

MCGILL, W. J. Neural counting mechanisms and energy detection in audition. *Journal of Mathematical Psychology*, 1967, **4**, 351–376.

MCGILL, W. J., & GOLDBERG, J. P. Pure-tone intensity discrimination and energy detection. *Journal of the Acoustical Society of America*, 1968, **44**, 576–581.

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

PARZEN, E. *Stochastic processes.* San Francisco: Holden-Day, 1962.

PERKEL, D. H., & BULLOCK, T. H. Neural coding. *Bulletin of the Neurosciences Research Program*, 1968, **6**, 221–348.

Rodieck, R. W., Kiang, N. Y-S., & Gerstein, G. L. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophysics Journal*, 1962, **2**, 351–368.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, 1967, **30**, 769–793.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. Some possible neural correlates of combination tones. *Journal of Neurophysiology*, 1969, **32**, 402–423.

Schneider, B., & Lane, H. Ratio scales, category scales, and variability in the production of loudness and softness. *Journal of the Acoustical Society of America*, 1963, **35**, 1953–1961.

Shiryaev, A. N. On optimum methods in quickest detection problems. *Theory of Probability and Its Applications*, 1963, **8**, 22–46.

Siebert, W. M. Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik*, 1965, **2**, 206–215.

Siebert, W. M. Stimulus transformations in the peripheral auditory system. In P. A. Kohlers & M. Eden (Eds.), *Recognizing patterns.* Cambridge, Mass.: M.I.T. Press, 1968.

Siebert, W. M. Frequency discrimination in the auditory system: Place or periodicity mechanisms? *Proceedings of the IEEE*, 1970, **58**, 723–730.

Stevens, J. C., & Hall, J. W. Brightness and loudness as functions of stimulus duration. *Perception and Psychophysics*, 1966, **1**, 319–327.

Stevens, J. C., & Stevens, S. S. Brightness function: Effects of adaptation. *Journal of the Optical Society of America*, 1963, **53**, 375–385.

Stevens, S. S. On the psychophysical law. *Psychological Review*, 1957, **64**, 153–181.

Stevens, S. S. Duration, luminance, and the brightness exponent. *Perception and Psychophysics*, 1966, **1**, 96–100.

Stevens, S. S. Neural events and the psychophysical law. *Science*, 1970, **170**, 1043–1050.

Stevens, S. S., & Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 1957, **54**, 377–411.

Stone, M. Models for choice-reaction time. *Psychometrika*, 1960, **25**, 251–260.

Swensson, R. G. The elusive tradeoff: Speed versus accuracy in a visual discrimination task. Unpublished doctoral dissertation, University of Michigan, 1969.

Swensson, R. G., & Edwards, W. Response strategies in a two-choice reaction task with a continuous cost for time. *Journal of Experimental Psychology*, 1971, **88**, 67–81.

Tippett, L. H. C. On the extreme individuals and the range of samples taken from a normal population. *Biometrika*, 1925, **17**, 364–387.

Vaughan, H. G., Costa, L. D., & Gilden, L. The functional relation of visual evoked response and reaction time to stimulus intensity. *Vision Research*, 1966, **6**, 645–656.

APPENDIX A

Proofs of Equations for Models with Strong Signals

Proof that $p(1 \,|\, s_1, s_2)$ *decreases with decreasing* k.

From the assumption

$$\frac{\mu_1}{\mu_2} > \frac{1 + \left(\dfrac{1}{2k-1}\right)^{\frac{1}{2}}}{1 - \left(\dfrac{1}{2k-1}\right)^{\frac{1}{2}}}$$

we conclude immediately that $\rho > \frac{1}{2}\left[1 + \left(\dfrac{1}{2k-1}\right)^{\frac{1}{2}}\right]$. Thus, for $x \geq \rho$,

$$\frac{\dfrac{(2k-1)!}{[(k-1)!]^2}\, x^{k-1}(1-x)^{k-1}}{\dfrac{(2k-3)!}{[(k-2)!]^2}\, x^{k-2}(1-x)^{k-2}} = 2\left(\frac{2k-1}{k-1}\right)x(1-x) < 1.$$

So

$$p_k(1\,|\,s_1,s_2) = 1 - \frac{(2k-1)!}{[(k-1)!]^2}\int_\rho^1 [x(1-x)]^{k-1}dx > 1 - \frac{(2k-3)!}{[(k-1)!]^2}\int_\rho^1 [x(1-x)]^{k-1}dx$$

$$= p_{k-1}(1\,|\,s_1,s_2).$$

Proof of Equation 9

Let f_i denote the density of $\mathbf{T}_i$, then

$$P(\mathbf{T}_2 - \mathbf{T}_1 \geq \epsilon) = \int_\epsilon^\infty \int_{\max(0,-t)}^\infty f_2(x+t)f_1(x)dxdt$$

$$= \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \int_\epsilon^\infty e^{-\mu_2 t} \int_{\max(0,-t)}^\infty [x(x+t)]^{k-1}e^{-(\mu_1+\mu_2)x}dxdt.$$

Observe that

$$\int_0^\infty [x(x+t)]^{k-1}e^{-(\mu_1+\mu_2)x}dx = \int_0^\infty x^{k-1} \sum_{i=0}^{k-1} \binom{k-1}{i}x^i t^{k-1-i}e^{-(\mu_1+\mu_2)x}dx$$

$$= \sum_{i=0}^{k-1} \binom{k-1}{i}t^{k-1-i}\frac{(k-1+i)!}{(\mu_1+\mu_2)^{k+i}}.$$

For $\epsilon \geq 0$, $t \geq 0$ and so $\max(0,-t) = 0$, and substituting and simplifying yields the first part of Equation 9. The second part, $\epsilon < 0$, requires separating the integral at 0,

$$P(\mathbf{T}_2 - \mathbf{T}_1 \geq \epsilon) = \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \left\{ \int_\epsilon^0 e^{-\mu_2 t} \int_{-t}^\infty [x(x+t)]^{k-1}e^{-(\mu_1+\mu_2)x}dxdt \right.$$

$$+ \int_0^\infty e^{-\mu_2 t} \int_0^\infty [x(x+t)]^{k-1}e^{-(\mu_1+\mu_2)x}dxdt$$

$$= \frac{(\mu_1\mu_2)^k}{[(k-1)!]^2} \int_\epsilon^0 e^{\mu_1 t} \sum_{i=0}^{k-1} \binom{k-1}{i}t^{k-1-i}\frac{(k+i-1)!}{(\mu_1+\mu_2)^{k+i}} dt + P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0)$$

$$= (1-\rho)^k \sum_{i=0}^{k-1} \binom{k+i-1}{i}\rho^k \left[1 - e^{\mu_1\epsilon} \sum_{j=0}^{k-1-i}\frac{(-\mu_1\epsilon)^j}{j!} \right.$$

$$\left. \left. + P(\mathbf{T}_2 - \mathbf{T}_1 \geq 0) \right] \right]$$

Proof of Equation 12

The proof of Equation 12 parallels that of Equation 5:

$$P\left(\frac{\mathbf{T}_2}{\mathbf{T}_1} \geq \delta\right) = \int_\delta^\infty P\left(\frac{\mathbf{T}_2}{\mathbf{T}_1} = t\right)dt$$

$$= \int_\delta^\infty \int_0^\infty f_2(ty)yf_1(y)dydt$$

$$= \frac{(2k-1)!}{[(k-1)!]^2} \int_\delta^\infty \frac{\left(\frac{\mu_2}{\mu_1}\right)^k t^{k-1}}{\left(1+\frac{\mu_2}{\mu_1}t\right)^{2k}} dt,$$

which yields Equation 12 by substituting

$$x = \frac{1}{1+\frac{\mu_2}{\mu_1}t}.$$

APPENDIX B

Proofs of Equations for Models with Weak Signals

Proofs of Equations 33–37

Assume $J = \kappa = 1$, a signal of duration T, and the response rule that the first IAT is compared with β and respond Y if it is $\leq \beta$; N otherwise. Let Y_i denote the time of the ith pulse. For $\beta \leq T$,

$$l_{sY}(t) = \frac{1}{p(Y\,|\,s)} \int_{\max(0,t-\beta)}^{t} P(Y_2 = t\,|\,Y_1 = x \& s)P(Y_1 = x\,|\,s)dx$$

$$= \frac{1}{p(Y\,|\,s)} \begin{cases} \int_0^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx, & t < \beta \\[2mm] \int_{t-\beta}^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx. & \beta \leq t < T \\[2mm] \int_{t-\beta}^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dt + e^{-\mu T}\int_T^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T \leq t < T+\beta \\[2mm] e^{-\mu T}\int_{t-\beta}^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(Y\,|\,s)} \begin{cases} \mu^2 t e^{-\mu t}, & t < \beta \\ \mu^2 \beta e^{-\mu t}, & \beta \leq t < T \\ \nu e^{-(\mu-\nu)T}e^{-\nu t}[(\mu-\nu)(T-t)+\mu\beta], & T \leq t < T+\beta \\ \nu^2 \beta e^{-(\mu-\nu)T}e^{-\nu t}, & T+\beta \leq t. \end{cases}$$

Similarly,

$$l_{sN}(t) = \frac{1}{p(Y\,|\,s)} \begin{cases} 0, & t < \beta \\ P(Y_1 = t-\beta\,|\,s)P(Y_2 - Y_1 \geq \beta\,|\,Y_1 = t-\beta \& s) & t \geq \beta \end{cases}$$

$$= \frac{1}{p(N\,|\,s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu(t-\beta)}e^{-\mu\beta}, & \beta \leq t < T \\ \mu e^{-\mu(t-\beta)}e^{-\mu(T-t+\beta)}e^{-\nu(t-T)}, & T \leq t < T+\beta \\ e^{-\mu T}\nu e^{-\nu(t-\beta-T)}e^{-\nu\beta}, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(N\,|\,s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu t}, & \beta \leq t < T \\ \mu e^{-(\mu-\nu)T}e^{-\nu t}, & T \leq t < T+\beta \\ \nu e^{-(\mu-\nu)T}e^{-\nu t}, & T+\beta \leq t. \end{cases}$$

The argument for $\beta > T$ is similar:

$$l_{sY}(t) = \frac{1}{p(Y\,|\,s)} \begin{cases} \int_0^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx, & t < T \\[2mm] \int_0^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dx + \int_T^t e^{-\mu T}\nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T \leq t < \beta \\[2mm] \int_{t-\beta}^T \mu e^{-\mu x}e^{-\mu(T-x)}\nu e^{-\nu(t-T)}dx + e^{-\mu T}\int_T^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & \beta \leq t < T+\beta \\[2mm] e^{-\mu t}\int_{t-\beta}^t \nu e^{-\nu(x-T)}\nu e^{-\nu(t-x)}dx, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(Y|s)} \begin{cases} \mu^2 t e^{-\mu t}, & t < T \\ \mu\nu e^{-(\mu-\nu)T} T e^{-\nu t} + \nu^2 e^{-(\mu-\nu)T}(t-T)e^{-\nu t}, & T \leq t < \beta \\ \mu\nu e^{-(\mu-\nu)T}(T-t+\beta)e^{-\nu t} + \nu^2 e^{-(\mu-\nu)T}(t-T)e^{-\nu t}, & \beta \leq t < T+\beta \\ \nu^2 \beta e^{-(\mu-\nu)T} e^{-\nu t}, & T+\beta \leq t. \end{cases}$$

$$l_{sN}(t) = \frac{1}{p(N|s)} \begin{cases} 0, & t < \beta \\ \mu e^{-\mu(t-\beta)} e^{-\mu(T-t+\beta)} e^{-\nu(t-T)}, & \beta \leq t < T+\beta \\ e^{-\mu T}\nu e^{-\nu(t-\beta-T)} e^{-\nu\beta}, & T+\beta \leq t \end{cases}$$

$$= \frac{1}{p(N|s)} e^{-(\mu-\nu)T} e^{-\nu t} \begin{cases} 0, & t < \beta \\ \mu, & \beta \leq t < T+\beta \\ \nu, & T+\beta \leq t. \end{cases}$$

Since these are densities, the expressions for $p(Y|s)$ and $p(Y|n)$ may be obtained by integrating them and setting that $= 1$. Equations 33 and 34 result. To obtain Equations 36 and 37 for MRTs, we first calculate the Fourier transforms:

$$p(Y|s)L_{sY}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right)^2 (1 - e^{(i\omega+\mu)\beta}) - \left(\frac{\mu}{i\omega+\mu}\right)\mu\beta e^{-(i\omega+\mu)T}$$

$$+ \left(\frac{\nu}{i\omega+\nu}\right) e^{-(i\omega+\mu)T}\left[\mu\beta - \left(\frac{\mu-\nu}{i\omega+\nu}\right)(1 - e^{-(i\omega+\nu)\beta})\right]$$

$$p(N|s)L_{sN}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right)(e^{-(i\omega+\mu)\beta} - e^{-(i\omega+\mu)T}) + \frac{e^{-(i\omega+\mu)T}}{i\omega+\nu}[\mu - (\mu-\nu)e^{-(i\omega+\nu)\beta}].$$

We then use the fact that

$$\int_0^\infty t f(t)dt = i \lim_{\omega \to 0} \frac{dF(\omega)}{d\omega}$$

to obtain the MRTs.

Proofs of Equations 38–40

Observe that no response is initiated later than T and, since the residual latency is bounded by τ, no response takes longer than $\tau + T$.

The calculation of Y responses is much as before:

$$l_{sY}(t) = \frac{1}{p(Y|s)} \begin{cases} \int_0^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx & t < \beta \\ \int_{t-\beta}^t \mu e^{-\mu x}\mu e^{-\mu(t-x)}dx, & \beta \leq t < T \\ 0 & T \leq t \end{cases}$$

$$= \frac{1}{p(Y|s)} \begin{cases} \mu^2 t e^{-\mu t}, & t < \beta \\ \mu^2 \beta e^{-\mu t}, & \beta \leq t < T \\ 0, & T \leq t. \end{cases}$$

The N responses arise in two ways: either the first IAT is wholly within $(0,T)$ and it exceeds β or there is at most one pulse in $(0,T)$:

$$l_{sN}(t) = \frac{1}{p(N|s)} \begin{cases} 0 & t < \beta \\ \mu e^{-\mu(t-\beta)}e^{-\mu\beta} = \mu e^{-\mu t} & \beta \leq t < T \\ (1 + \mu\beta)e^{-\mu T} & t = T \\ 0 & T < t. \end{cases}$$

Note that the expressions for n trials are obtained by replacing μ by ν.

By integrating either density (the latter is easier), we obtain Equation 38. The Fourier transforms are:

$$p(\mathrm{Y}\,|\,s)L_{s\mathrm{Y}}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right)^2 (1 - e^{-(i\omega+\mu)\beta}) - \left(\frac{\mu^2\beta}{i\omega+\mu}\right)e^{-(i\omega+\mu)T}$$

$$p(\mathrm{N}\,|\,s)L_{s\mathrm{N}}(\omega) = \left(\frac{\mu}{i\omega+\mu}\right) (e^{-(i\omega+\mu)\beta} - e^{-(i\omega+\mu)T}) + (1 + \mu\beta)e^{-(i\omega+\mu)T},$$

from which Equations 39 and 40 follow.

APPENDIX C

Proofs of Equations for Simple Reaction Times

Given the definitions of l and l_0, Equations 43 and 44, and the independent exponential distribution of IATs, it is not difficult to see that they satisfy the following pair of functional-integral equations:

$$l(x,t) = \begin{cases} \int_0^t \nu e^{-\nu y} l_0(x - y, t - y)dy, & t < x \\[2mm] \int_0^x \nu e^{-\nu y} l_0(x - y, t - y)dy + e^{-\nu x}\int_x^t \mu e^{-\mu(y-x)} l_0(0, t - y)dy, & t \geq x \end{cases} \qquad [\text{C1}]$$

$$l_0(x,t) = \begin{cases} \nu e^{-\nu t}, & t < x \\ e^{-\nu x}\mu e^{-\mu(t-x)} & x \leq t \leq \beta \\ e^{-\nu x}e^{-\mu(\beta-x)}l(0, t - \beta), & x < \beta \leq t \end{cases} \Bigg\} \, x < \beta \\[2mm] \begin{cases} \nu e^{-\nu t} & t \leq \beta \\ e^{-\nu\beta}l(x - \beta, t - \beta), & \beta < t \end{cases} \Bigg\} \, x \geq \beta. \qquad [\text{C2}]$$

Setting $x = 0$ in Equation C2 yields Equation 45.
For $x \leq \beta$, t, Equation C2 yields

$$l_0(x,t) = \begin{cases} e^{-\nu x}\mu e^{-\mu(t-x)}, & x \leq t \leq \beta \\ e^{-\mu x}e^{-\mu(\beta-x)}l(0, t - \beta), & x \leq \beta < t \end{cases}$$

$$= e^{-\nu x}e^{\mu x}\begin{cases} \mu e^{-\mu t} \\ e^{-\mu\beta}l(0, t - \beta) \end{cases}$$

$$= e^{(\mu-\nu)x}l_0(0,t).$$

Using this in Equation C1 yields for $x \leq \beta$,

$$l(x,t) = \begin{cases} \nu^2 t e^{-\nu t}, & t < x \\[2mm] e^{(\mu-\nu)x}\left[\nu\int_0^x e^{-\mu y}l_0(0, t - y)dy + \mu\int_x^t e^{-\mu y}l_0(0, t - y)dy\right], & t \geq x. \end{cases} \qquad [\text{C3}]$$

Proof of Equation 46

Define

$$\varphi(t) = l_0(0,t)e^{\mu t}/\mu.$$

By Equation 45, φ satisfies

$$\varphi(t) = \begin{cases} 1, & 0 \leq t < \beta \\ \mu \int_0^{t-\beta} \varphi(y)dy, & \beta \leq t. \end{cases}$$

We show that

$$\varphi(t) = \sum_{i=0}^{m-1} a(m-1-i)\frac{\mu^i[t-(m-1)\beta]^i}{i!}, \quad (m-1)\beta \leq t < m\beta,$$

solves this equation if a is defined recursively as in Equation 47. Since this function is 1 for $0 \leq t < \beta$, we consider $t \geq \beta$. Suppose m is such that $(m-1)\beta \leq t < m\beta$, then by induction

$$\mu \int_0^{t-\beta} \varphi(z)dz = \mu \left[\sum_{i=1}^{m-2} \int_{(i-1)\beta}^{i\beta} \varphi(z)dz + \int_{(m-2)\beta}^{t-\beta} \varphi(z)dz \right]$$

$$= \mu \left\{ \sum_{i=1}^{m-2} \int_0^{\beta} \varphi[y+(i-1)\beta]dy + \int_0^{t-(m-1)\beta} \varphi[y+(m-2)\beta]dy \right\}$$

$$= \mu \left\{ \sum_{i=1}^{m-2} \int_0^{\beta} \sum_{j=0}^{L-1} a(i-1-j)\frac{\mu^j y^j}{j!}dy + \int_0^{t-(m-1)\beta} \sum_{j=0}^{m-2} a(m-2-j)\frac{\mu^j y^j}{j!}dy \right\}$$

$$= \sum_{i=1}^{m-2} \sum_{k=1}^{i} a(i-k)\frac{(\mu\beta)^k}{k!} + \sum_{k=1}^{m-1} a(m-1-k)\frac{\mu^k[t-(m-1)\beta]^k}{k!}$$

$$= a(m-1) + \sum_{k=1}^{m-1} a(m-1-k)\frac{\mu^k[t-(m-1)\beta]^k}{k!}$$

$$= \varphi(t).$$

Another approach to solving integral equations of retarded type is to differentiate them and attempt to reduce the problem to one of differential equations of retarded type. These have been studied in some detail (see Bellman & Cooke, 1963; El'sgol'ts, 1966).

Care must be taken, however, since information is lost. For example, our integral equation for φ says φ is discontinuous at $t = \beta$. If we differentiate, ignoring all potential discontinuities,

$$\varphi'(t) = \mu\varphi(t-\beta), \quad t > \beta,$$
$$\varphi(t) = 1, \quad 0 \leq t < \beta.$$

This is dealt with on page 6 on El'sgol'ts and the solution is

$$\varphi(t) = \sum_{i=0}^{m-1} \mu^i(t-i\beta)^i/i!, \quad (m-1)\beta \leq t < m\beta.$$

It is easy to see that this does not solve the integral equation because it is continuous at $t = \beta$.

Proof of Equation 48

Rewrite Equation 45 and use the linear term of the Taylor series,

$$e^{\mu t}l_0(0,t) = \mu \begin{cases} 1, & t \leq \beta \\ \int_0^{t-\beta} e^{\mu v}l_0(0,y)dy & t > \beta \end{cases}$$

$$\cong \mu \begin{cases} 1 \\ \int_0^{t} e^{\mu v}l_0(0,y)dy - \beta e^{\mu t}l_0(0,t). \end{cases}$$

Differentiate this for $t > \beta$ and collect terms

$$l'_0(0,t) = -\mu' l_0(0,t),$$

where

$$\mu' = \frac{\mu^2 \beta}{1 + \mu\beta}.$$

Equation 48 follows by integration except for a constant of integration which is evaluated by setting $\int_0^\infty l_0(0,t)dt = 1$.

Proof of Equation 50

From Equation 45,

$$
\begin{aligned}
L_0(0,\omega) &= \int_0^\infty e^{-i\omega t} l_0(0,t)dt \\
&= \int_0^\beta e^{-i\omega t}\mu e^{-\mu t}dt + \int_\beta^\infty e^{-i\omega t}\mu e^{-\mu t}\int_0^{t-\beta} e^{\mu y} l_0(0,y)dy \\
&= \frac{\mu}{i\omega + \mu}[1 - e^{-(i\omega+\mu)\beta}] + \mu e^{(i\omega+\mu)\beta}\int_0^\infty e^{-(i\omega+\mu)z}\int_0^z e^{\mu y} l_0(0,y)dy.
\end{aligned}
$$

Integrating by parts, taking into account that $l_0(0,t) \to 0$ as $t \to \infty$ because l_0 is a density, we obtain

$$L_0(0,\omega) = \frac{\mu}{i\omega + \mu}[1 - e^{-(i\omega+\mu)\beta} + e^{-(i\omega+\mu)\beta}L_0(0,\omega)].$$

Solving yields Equation 50.

Proof of Equation 51

By Equation C3,

$$
\begin{aligned}
L(x,\omega) &= \int_0^x e^{-i\omega t} l(x,t)dt + \int_x^\infty e^{-i\omega t} l(x,t)dt \\
&= \int_0^x e^{-i\omega t}\nu^2 t e^{-\nu t}dt + e^{(\mu-\nu)x}\int_x^\infty e^{-i\omega t}\nu\int_0^x e^{-\mu y} l_0(0, t - y)dydt \\
&\qquad\qquad + e^{(\mu-\nu)x}\int_x^\infty e^{-i\omega t}\mu\int_x^t e^{-\mu y} l_0(0, t - y)dydt.
\end{aligned}
$$

The first integral is trivial. The second one can be dealt with by a change of variable and an integration by parts:

$$
\begin{aligned}
\nu e^{(\mu-\nu)x}\int_x^\infty e^{-(i\omega+\mu)t}\int_{t-x}^t e^{\mu z} l_0(0,z)dzdt &= \frac{\nu e^{(\mu-\nu)x}}{i\omega + \mu}\left\{ - e^{-(i\omega+\mu)t}\int_{t-x}^t e^{\mu z} l_0(0, z)dz\bigg|_x^\infty \right. \\
&\qquad \left. + \int_x^\infty e^{-(i\omega+\mu)t}[e^{\mu t} l_0(0,t) - e^{\mu(t-x)} l_0(0, t - x)]dt\right\} \\
&= \frac{\nu e^{(\mu-\nu)x}}{i\omega + \mu}\left\{ \mu x e^{-(i\omega+\mu)x} \right. \\
&\qquad \left. + \left[L_0(0,\omega) - \frac{\mu}{i\omega + \mu}\right][1 - e^{-(i\omega+\mu)x}]\right\}.
\end{aligned}
$$

A similar treatment of the third integral yields

$$\frac{\mu e^{(\mu-\nu)x}e^{-(i\omega+\mu)x}}{i\omega + \mu}L_0(0,\omega).$$

Adding the three terms gives Equation 51.

Proof of Equation 54

In the following estimates of $l(x,t)$ we are only concerned with $t \geqslant x$ and so it will prove convenient to write the second argument as $t + x, t \geqslant 0$. Consider, first, the case $x \leqslant \beta$. The plan is to substitute the approximate solution for $l_0(0,t)$, Equation 48, into the second part of Equation C3. Observe that if we confine ourselves to $t > \beta$, then $t + x - y \leqslant \beta$ iff $y \geqslant t + x - \beta > x$, hence the two integrals of Equation C3 divide into three:

$$l(x, t + x) \cong e^{(\mu-\nu)x}\left\{\nu \int_0^x e^{-\mu\nu}\mu'e^{-\mu'/\mu}e^{-\mu'(t+x-\nu)}dy + \mu \int_x^{t+x-\beta} e^{-\mu\nu}\mu'e^{-\mu'/\mu}e^{-\mu'(t+x-\nu)}dy \right.$$

$$\left. + \mu \int_{t+x-\beta}^{t+x} e^{-\mu\nu}\mu e^{-\mu(t+x-\nu)}dy\right\}.$$

Carrying out the integrations and taking note of the fact that $\mu = \mu' + \mu/(1 + \mu\beta)$ yields

$$l(x, x + t) \cong A''e^{-\mu t} + B''e^{-\mu't}, \quad x \leqslant \beta < t,$$

where A'', B'' are functions of ν, μ, β, and x which we do not need to specify.

For $x > \beta$, we proceed by a somewhat different argument.[12] Let $\mathbf{Y}$ be the time of the last pulse, if any, prior to the signal onset at time x, or 0 otherwise, and let $\mathbf{Y}'$ be the time of the first pulse after the signal onset. Since, in a Poisson process, the time from x back to the preceding pulse is exponentially distributed with the parameter of the process, we see that

$$P(\mathbf{Y}' = y, \mathbf{Y}' - \mathbf{Y} > \beta | \mathbf{S} = x > \beta) = \mu e^{-\mu\nu}\begin{cases} e^{-\nu(\beta-\nu)}, & y \leqslant \beta \\ 1, & y > \beta. \end{cases}$$

Observe that if $t > 2\beta$, then $t - y \leqslant \beta$ iff $y \geqslant t - \beta > \beta$, so

$$l(x, t + x) = \int_0^t P(\mathbf{Y}' = y, \mathbf{Y}' - \mathbf{Y} > \beta | \mathbf{S} = x > \beta)l_0(0, t - y)dy$$

$$\cong \int_0^\beta \mu e^{-\mu\nu}e^{-\nu(\beta-\nu)}\mu'e^{-\mu'/\mu}e^{-\mu'(t-\nu)}dy + \int_\beta^{t-\beta} \mu e^{-\mu\nu}\mu'e^{-\mu'/\mu}e^{-\mu'(t-\nu)}dy$$

$$+ \int_{t-\beta}^t \mu e^{-\mu\nu}\mu e^{-\mu(t-\nu)}dy$$

$$= A''e^{-\mu t} + B''e^{-\mu't}, \quad x > \beta, \quad t > 2\beta,$$

where A'', B'' are functions of ν, μ, and β in the range $x > \beta$.

Thus, the density of RTs $\geqslant \tau$ to a signal with onset x is

$$f(x, t + x) = \int_0^\tau l(x, t + x - y)r(y)dy, \quad t \geqslant \tau.$$

For $t > \tau + 2\beta$, $t - y > 2\beta$ and so we may substitute the above approximations for $l(x, t + x - y)$, yielding

$$f(x, t + x) = A'e^{-\mu t} + B'e^{-\mu't}, \quad \text{for} \quad t > \tau + 2\beta,$$

----

[12] We are indebted to E. Thomas for pointing out this argument.

where A', B' are functions of ν, μ, β, x, and r. Assuming an arbitrary density g of foreperiods,

$$f_{R-S}(t) = C \int_0^\infty g(x) f(x, t + x) dx$$

$$= A e^{-\mu t} + B e^{-\mu' t}, \quad \text{for} \quad t > \tau + 2\beta,$$

where A, B are functions of ν, μ, β, g, and r.

Proof of Equation 55

First, we show that for $t \leq x$, $l(x,t) = l(t,t)$. By Equation C1, we see that it is sufficient to show that $l_0(x,t) = l_0(t,t)$. By Equation C3, this is true for $x \leq \beta$ and for $t \leq \beta < x$. For $\beta < t \leq x$, it is true provided that $l(x - \beta, t - \beta) = l(t - \beta, t - \beta)$. By a finite induction of the form $x - i\beta$, the proof is reduced to the already proved case of $x \leq \beta$. Now, with $t \leq x$, Equations C1 and C2 yield

$$l(t,t) = \int_0^t \nu e^{-\nu y} l_0(t - y, t - y) dy$$

$$= \begin{cases} \displaystyle\int_0^t \nu e^{-\nu y} \nu e^{-(t-y)} dy, & t \leq \beta \\ \displaystyle\int_0^\beta \nu e^{-\nu y} \nu e^{-\nu(t-y)} dy + \int_\beta^t \nu e^{-\nu y} e^{-\nu\beta} l(t - y - \beta, t - y - \beta) dy, & t > \beta \end{cases}$$

$$= \begin{cases} \nu^2 t e^{-\nu t}, & t \leq \beta \\ \nu^2 \beta e^{-\nu t} + \nu e^{-\nu t} \displaystyle\int_0^{t-\beta} e^{\nu z} l(z,z) dz, & t > \beta. \end{cases}$$

Proof of Equation 59

Assume $t \geq \tau' + \tau$. Then

$$f_R(t) = K' \int_t^\infty \lambda e^{-\lambda x} \int_0^t l(x, t - y) r(y) dy dx$$

$$= K' \int_t^\infty \lambda e^{-\lambda x} dx \int_0^\tau l(t - y, t - y) r(y) dy$$

$$= K' e^{-\lambda t} \int_0^\tau D e^{-\nu'(t-y)} r(y) dy$$

$$= K' D e^{-(\lambda+\nu')t} \int_0^\tau e^{\nu' y} r(y) dy.$$

Partial Results about $l(x,t)$ *for* $x \geq \beta$

From the argument given in deriving Equation 54, for $x > \beta$,

$$l(x, t + x) = \begin{cases} \mu e^{-\nu\beta} e^{-(\mu-\nu)t} \displaystyle\int_0^t e^{(\mu-\nu)y} l_0(0,y) dy, & 0 \leq t \leq \beta \\ \mu e^{-\nu\beta} e^{-(\mu-\nu)t} \displaystyle\int_{t-\beta}^t e^{(\mu-\nu)y} l_0(0,y) dy + \mu e^{-\mu t} \int_0^{t-\beta} e^{\mu y} l_0(0,y) dy, & t > \beta. \end{cases}$$

By routine integration,

$$\int_0^\infty e^{-i\omega t} l(x, t + x) dt = \frac{\mu e^{-\nu \beta} L_0(0,\omega)}{\mu + i\omega},$$

where $L_0(0,\omega)$ is given by Equation 50. So, using Equation 55,

$$L(x,\omega) = \int_0^\infty e^{-i\omega t} l(x,t) dt$$

$$= \int_0^x e^{-i\omega t} l(t,t) dt + e^{-i\omega x} \int_0^\infty e^{-i\omega \nu} l(x, y + x) dy$$

$$= \left(\frac{\nu}{\nu + i\omega}\right)^2 (1 - [(\nu + i\omega)\beta + 1] e^{-(\nu + i\omega)\beta}) + \int_\beta^x e^{-i\omega t} l(t,t) dt + \frac{e^{-i\omega x} \mu e^{-\nu \beta} L_0(0,\omega)}{\mu + i\omega}.$$

The problem in going further is how to evaluate the integral in the second term. One possibility is to use the approximation of Equation 58; however, we are uncertain how good it is near β. If so, then it is easy to calculate the mean decision time for $x > \beta$, which together with Equation 52 allows us to calculate the MRT for any distribution of foreperiods. As this expression is both complex and approximate, we do not present it.