

# Evidence from auditory simple reaction times for both change and level detectors

STEPHEN L. BURBECK

*Linus Pauling Institute of Science and Medicine, Palo Alto, California*

and

R. DUNCAN LUCE

*Harvard University, Cambridge, Massachusetts*

We examine the form of distributions of simple reaction time. The stimuli we use are the offset of weak pure tones masked by wide-band noise. The hazard functions of the RT distributions (i.e., the probability of a response given that one has not already occurred) are monotone increasing for very weak tones but become peaked for stronger tones. Of the models available in the literature, none is very satisfactory, although two can account for the general qualitative shape of the peaked hazard functions. We propose a model wherein both a change and a level detector function in parallel. If one assumes that the change detector and level detector have slightly different thresholds, this model can account for both the monotone increasing and the peaked hazard functions.

This paper reports new psychophysical evidence that there exist two quite distinct classes of mechanisms for detecting changes in the intensity of auditory signals. One type, called a change detector, is sensitive to abrupt changes in the signal and responds transiently to such changes. The other type, called a level detector, is sensitive to the absolute level of the signal. The level detector may be a bit slow to respond to change, but, since it tracks the level of the signal, it remains capable of reporting the changed signal intensity long after the change is completed. The change detector responds quickly to a change but is less persistent in that, once the change has receded sufficiently into the past, it is unlikely to initiate a response. This difference in persistence is what we attempt to exploit: it implies that the two types of detector should produce quite different forms of simple reaction-time distribution.

Physiological evidence for both classes of detectors exists in vision and in audition (Cleland et al., 1971; Enroth-Cugel & Robson, 1966; Gerstein, Butler, & Erulkar, 1968; Marrocco, 1976; Møller, 1969; Pfeiffer, 1966). Psychophysical evidence is less plentiful, particularly in audition. Macmillan (1971, 1973) reports auditory detection and recognition data that seem to

The research reported here is part of S. L. Burbeck's PhD dissertation, University of California at Irvine. The data were collected at the Psychophysical Laboratory, Harvard University. It was supported in part by grants from the National Science Foundation to the University of California at Irvine and to Harvard University, R. D. Luce, principal investigator. Reprint requests should be sent to: Stephen L. Burbeck, Linus Pauling Research Institute, 440 Page Mill Road, Palo Alto, California 94306.

implicate both change and level detectors, as do the results of Ahumada, Marken, and Sandusky (1975). And Tolhurst (1975) reports clear evidence of both change and level detection in visual simple reaction time.

We infer the existence of the two classes of detector from qualitative features of simple reaction-time distributions that are obtained with a response-terminated random foreperiod paradigm (Luce & Green, 1970). These qualitative features are described in terms of a transform of the distribution called the *hazard function*. Thus, we must introduce the hazard function and some of its properties before we can discuss the qualitative features that implicate change and level detectors.

## Transforming Reaction-Time Distributions

The behavior of change and level detectors may differ in many respects, but the difference that is manifested in simple reaction time is one of persistence: the change detector, if it fails to respond to a change, is thereafter unlikely to respond because those characteristics of its input which indicate a change do not persist, whereas the level detector is unhindered by the passage of time because successive estimates of the level remain informative no matter how much time has passed after the change. These differences in persistence are difficult to discern in the reaction-time distribution function or density function, but are directly reflected in a transform of the distribution function called the hazard function or the conditional density function.

Let  $T$  be a nonnegative random variable, such as the latency of a detection mechanism, and let  $f$  be its

density and  $F$  be its distribution function, that is,

$$F(t) = \int_0^t f(x) dx.$$

The hazard function of  $f$  is the density of an observation occurring at time  $t$  given that it has not occurred earlier, that is,

$$h(t) = \frac{f(t)}{1 - F(t)} = -\frac{d}{dt} \ln[1 - F(t)]. \quad (1)$$

By integrating Equation 1, solving for  $F(t)$ , and differentiating the result, one obtains the inverse relation

$$f(t) = h(t) \exp \left[ -\int_0^t h(x) dx \right]. \quad (2)$$

Thus, when one thinks of a probabilistic process as unfolding in time and one is primarily interested in the termination of that process (e.g., the termination of a detection process), one is really talking about the hazard function rather than the density function because a process can terminate in the future only if it has not done so already.

The quantity  $\ln[1 - F(t)]$  in Equation 1 is called the *log survivor function*. Reaction-time data have sometimes been plotted in the form of the log survivor function (Carterette et al., 1965; McGill, 1963; McGill & Gibbon, 1965; Snodgrass, 1969; Shannon, Note 1). With such plots, rough estimates of the hazard function can be obtained by eye, since the hazard function is the negative of the slope of the log survivor function.

The assumptions that make the hazard function useful for understanding simple reaction-time distributions are: (1) The sensory information is encoded as a stochastic process. This assumption is in accord with what we know about the physiology of the early stages of the auditory system (Kiang et al., 1965; Rose et al., 1967). (2) The decision process is an ongoing evaluation of the stochastic input—either continuously or in discrete steps—which responds when the probability of a false alarm first falls below a fixed criterion. That is, the evaluation of the input is equivalent to the computation of some possibly quite complex statistic from the stochastic input; the detector responds when this statistic first crosses a fixed criterion. This assumption is most supportable in the context of a random foreperiod paradigm, in which the stimulus context forces an ongoing decision process. Given these two assumptions, the hazard rate of the reaction-time distribution (if we ignore the influence of residual latencies) is a measure of the average accumulated information as a function of time.

In interpreting the data to be presented subsequently, we postulate that both a level detector and a change detector are independently at work—in parallel—so that either can produce a detection response. The hazard function of such a composite process turns out to be simply related to the hazard functions of each of its parts. Suppose the latencies of the change and level detectors,  $T_c$  and  $T_l$ , are independent random variables with hazard functions  $h_c$  and  $h_l$ , respectively. If the overall parallel system responds as soon as either detector is triggered, then the system latency is  $T = \min(T_c, T_l)$  and its hazard function is simply the sum

$$h(t) = h_c(t) + h_l(t). \quad (3)$$

The assumption that  $T_c$  and  $T_l$  are independent is untested and quite likely untestable. However, if the two processes receive inputs from distinct sets of neurons, the assumption is quite plausible. If not, then the assumption rests on further assumptions about the processing of those inputs. Because the neural code is stochastic, the decision processes can be thought of as depending upon some statistic computed from the input. In general, these statistics may either be independent (e.g., the mean and variance of the normal) or not (e.g., the mean and variance of the Poisson). Therefore, until change and level detectors can be isolated physiologically, the assumption of independence must remain unjustified.

### Qualitative Features of Change and Level Detectors

There is no need for us now to develop detailed models of change and level detectors, although elsewhere (Burbeck, 1979) a physiologically motivated model for a change detector is presented and tested against reaction-time data and Luce and Green (1972) have studied a simple level detector. Here it suffices to state the major qualitative distinctions between the two classes of devices under the general assumptions mentioned above together with two further assumptions. The first is that the signals are response terminated. The second is that the memory available to each detector is not unlimited. The assumption that the signals are response terminated (i.e., only one change in the intensity occurs prior to the subject's response) is not strictly necessary but simplifies the discussion of change detectors. The importance of the second assumption will become clear as we proceed.

Consider a signal that changes, say, decreases in intensity, abruptly at time  $t_0$ . The subject's task is to detect the change. We assume that during the foreperiod (prior to  $t_0$ ), when the signal is unchanging, any detector of the class we are considering here has some constant, presumably small, tendency to ter-

minate (i.e., produce a false alarm) as a result of random fluctuations in the neurally encoded signal. Hence, during the foreperiod, both  $h_c$  and  $h_l$  are constants. The major feature of a change detector is that shortly after  $t_0$ ,  $h_c$  increases markedly—how much, depends upon the magnitude of the change—and then, as the change recedes into the past, decreases again to a low value associated with the steady intensity of the signal after  $t_0$  (see Figure 1). The hazard rate increases as the ongoing decision process receives evidence of the change, although that evidence may at first be obscured by the stochastic nature of the input. The subsequent reduction in the hazard rate results from the memory limitation. Eventually the evidence of the change is “forgotten.” (If the memory were unlimited, then one could construct a model change detector that does not become less effective, because it can reexamine the entire history of its input.) By contrast, the hazard rate of a level detector is low up to  $t_0$ , then rises monotonically to

a higher value as neurally encoded evidence of the new level collects, and persists at that value until the intensity of the signal is changed again. In this case, the memory limitation does not result in a decreasing tail, because newly arriving information continues to indicate that the offset has occurred.

The hazard function of the composite detection process is the sum of the above two functions. If the peak in  $h_c$  is sufficiently large relative to  $h_l$ , then the composite hazard rate is low prior to  $t_0$ , rises to a peak, and then declines to an asymptotic value that is due largely to the level detector. If the peak in  $h_c$  is sufficiently small relative to  $h_l$  (perhaps because the change is nearer the threshold of the change detector than that of the level detector), the composite hazard function may be monotone increasing. Schematic plots of change and level detector hazard functions are shown in Figure 1 together with their sum (Equation 3) and the corresponding log survivor function.

## THE EXPERIMENT

The empirical reaction-time distributions available in the literature are seldom based on enough trials to see the shape of the tail of the log survivor function. Green and Luce (1971) used large samples, and several of their distributions, although plotted in another way, had peaked hazard functions. Even clearer examples can be found in the data of Snodgrass (1969), although she did not use random foreperiods and her signals were not response-terminated. The most pronounced curvature we know of is the unpublished data of Shannon (Note 1), which was what led us to undertake this study.

The simple reaction-time paradigm we report here used weak, response-terminated signals with exponentially distributed random foreperiods (Carterette, Friedman, & Cosmides, 1965; Green & Luce, 1971; Luce & Green, 1970; Shannon, Note 1). We used weak signals for two reasons: (1) Reaction time to strong signals is rapid and most probably dominated by residual factors extraneous to the detection process, whereas with weak signals the reaction times are much slower, so the residual delays are less bothersome; and (2) the differences in persistence of the two sorts of detector do not manifest themselves until the signal has gone unnoticed for a few hundred milliseconds, which is an exceedingly rare occurrence unless one uses weak signals. We avoided the usual short-duration signal in favor of a response-terminated signal because analysis of the data was simplified if we ensured that, once the signal changed, it did not change again. We used exponentially distributed random foreperiods to eliminate the value of strategic anticipation responses, inasmuch as the exponential provided no information as to when the signal would appear (Green & Luce, 1971; Nickerson, 1967).

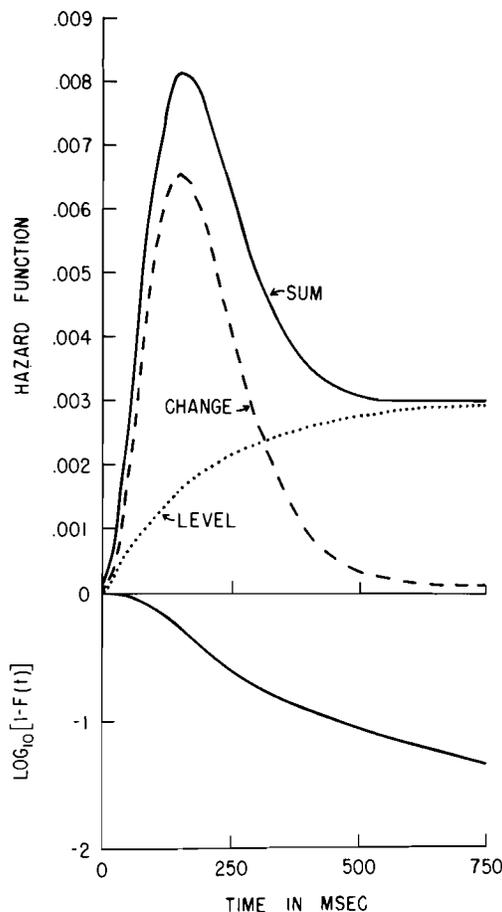


Figure 1. Schematic relations of signal intensity, the detector hazard functions, their sum, and the corresponding log survivor function. The level detector hazard function satisfies  $.003(1 - e^{-t/200})$  and the change detector  $.0385(t/100)^2e^{-t/50} + .0001$ .

The stimulus used in the first phase of the experiment was the offset of a weak pure-tone signal in broadband noise. We selected offsets because Shannon's data exhibited the clearest evidence for peaked hazard functions and this may have been due to his use of offsets rather than onsets. The broadband masker serves to reduce threshold differences between subjects and therefore permits the subjects to be run at the same time using the same signal parameters.

We gathered data at a range of frequencies (250, 1,000, and 4,000 Hz) because of the widespread belief that sensory encoding may be different at low frequencies from what it is at high frequencies: the periodicity of neural firing may be the critical variable at low frequencies, and the place of maximal stimulation of the basilar membrane may be critical for high frequencies (see Small, 1970). Although both types of encoding may be involved at 1,000 Hz, that frequency is commonly used and hence the resulting data are most comparable to those of other experimenters. Two intensities were used at each frequency. The intensities differed by only 2 dB because mean reaction time is exquisitely sensitive to intensity in the low range and we wished to see if other properties of the distribution were as sensitive.

As will become clear in a later section, we also required distributions that approximate the subject's minimum (i.e., residual) reaction times. The stimulus we chose was the offset of a loud, wide-band noise. There were two reasons for using noise rather than a tone. First, some evidence suggests that the internal effect of noise decays more rapidly than does that of a pure tone (Miller, 1948). Second, and perhaps more importantly, noise stimulates the entire basilar membrane and more of the auditory nerve than does a pure tone. Thus, the decision about change can, in principle, be made on the basis of whichever frequency-specific channel responds most quickly on each trial. In that case, the decision latency is the minimum of a possibly large number of parallel decision processes.

One potential drawback to using noise is that the total power of the signal is limited by consideration for the observers, and so the noise power available to each channel is less than would be available if a pure tone were used. Whether the tradeoff favors noise or a pure tone was not studied in depth, but an informal pilot experiment showed mean reaction times to be at least as short for noise as for a pure signal tone.

### Method

Three male observers (S.B., D.L., P.G.), including the first author, participated in the experiments (except that D.L. was unavailable when the strong-signal condition was run and P.G. was unavailable for one of the weak-signal conditions). They all had normal hearing. Other than S.B., they were paid at an hourly rate of \$2.25 plus a bonus, described below.

Each observer was seated in a separate sound-treated booth, and all received the same stimuli presented binaurally through

TDH-39 earphones driven in phase. The observers responded by pressing a microswitch. Except for the strong-signal case, masking noise was present continuously; its spectrum was flat (prior to the earphones which began to roll off at about 8 Hz), with a spectral density of 10 dB SPL. To avoid detectable off-frequency energy splatter from the abrupt offset, the signal channel was filtered with a bandwidth of 100 Hz centered on the signal frequency; hence, the decay lasted about 10 msec.

The signals run were: 250 Hz at  $P/N_0 = 20$ , 22 dB; 1,000 Hz at  $P/N_0 = 20$ , 22 dB; 4,000 Hz at  $P/N_0 = 24$ , 26 dB; and wide-band (0 to 10 kHz) noise at 70 dB SPL.

The temporal structure of a trial was as follows. The warning signal consisted of three successive 200-msec warning lights. Two hundred milliseconds after the offset of the third light, the foreperiod began. It was exponentially distributed with a mean of 4 sec in the 1,000-Hz,  $P/N_0 = 22$  dB case and of 3 sec in all others. (After collecting the initial data, it was decided that a mean wait of 4 sec was uncomfortably long.) The instructions to the observers included an explanation of the statistical properties of the exponential foreperiod distribution. In particular, it was made clear that the hazard function is constant and that "guessing" strategies are inherently pointless. Since the exponential is unbounded, we elected to truncate it at a duration of five times the mean, a cutoff affecting less than 1% of the trials. The tone was turned off at the end of the foreperiod and remained off until the last observer responded or until 3 sec had elapsed (an extremely rare occurrence). Then the tone was restored and feedback was provided as described below. After the completion of each block of trials, mean reaction times and the number of false alarms were printed by the computer (a PDP-15) and made available to the observers.

Trials were grouped into blocks of 75 to 100 trials, depending on conditions that affected the average length of each trial. A block typically lasted about 10 min, and a few minutes of rest were allowed between blocks. From five to nine blocks were run on each day. At most, the observers participated for 2 h per day. There were 6 days of training (2,855 trials) before the data reported here were collected. During this time, the observer's performance improved and stabilized.

To minimize unstable observer performance, trial-by-trial feedback was provided. Observers were rewarded for fast responses and penalized for false alarms (defined, for this purpose, as both anticipations and responses within the first 100 msec after signal offset). For the 1,000-Hz,  $P/N_0 = 22$  dB case, a single deadline was set for each observer (S.B., 400 msec; D.L., 430 msec; P.G., 460 msec), so that all three were about equally able to beat the deadline. The observers were informed each time they had beaten the deadline (by a flashing light) and were credited with three payoff points. They were also informed of each false alarm and debited one payoff point. For the remaining conditions, more detailed feedback was provided by a more elaborate deadline structure. Five bands were defined (again, according to the average speed of the individual observer). Responses after the longest deadline were not rewarded. The remaining four deadlines were spaced 50 msec apart, and each was worth an additional point. At the end of the 2-h session, \$2 was apportioned between D.L. and P.G. in proportion to their accumulated points for that session.

## DATA ANALYSIS

### Estimating Hazard Functions

The eye can read the hazard function roughly from plots of  $\ln[1 - F(t)]$  vs.  $t$ , but for more careful study of the detection process, better estimates are needed. Moreover, since we are making no distribution assumptions, we need distribution-free estimation procedures. Watson and Leadbetter (1964a, 1964b) discuss a variety of such methods, beginning with a

slightly formalized faired curve to the log survivor function and ending with rather more refined techniques.

Most estimation methods produce very ragged estimates of hazard functions, particularly in the tails (Rice & Rosenblatt, 1976). Roughly, they all depend upon having a smoothed sample density function obtained by linearly filtering the ordered events, that is, smoothing each event with the same smoothing function. The various estimation methods differ in the manner in which the smoothing function is applied, but all use the same width throughout the distribution. As a result, the estimate is ragged wherever the data are sparse, most notably in the tails of the distributions. Furthermore, while they are usually asymptotically unbiased estimators, they converge with the square root or, in some cases, with the cube root of the sample size.

A distribution-free estimation method recently proposed by Miller and Singpurwalla (1977) seems to offer substantial improvement over previous methods. They abandoned the linear smoothing for what they call random smoothing. This amounts to adjusting the width of the smoothing function to the density of events. Where events are closely packed, the smoothing function is narrow, and where they are sparse, the smoothing function is wide. The hazard function is estimated by a step function,  $H(i)$ , which is constant over each interval  $Z(i)$ ,  $Z(i-1)$ , where  $Z(i)$ ,  $i = 1, \dots, n$  is the order statistic from a sample of  $n$  reaction times, that is,  $Z(i)$  is the  $i^{\text{th}}$  smallest observed reaction time, and  $Z(0) = 0$ .  $H(i)$  is calculated on the basis of  $j$  consecutive ordered reaction times (where  $j$  is a parameter) as follows. Let

$$S(i) = (n - i + 1)[Z(i) - Z(i - 1)]. \quad (4)$$

In the present application,  $S(i)$  is just the difference between the  $i^{\text{th}}$  ordered reaction time and the one preceding it times the number of trials on which the response was longer than the  $i^{\text{th}}$  RT. The estimate of the hazard rate for the  $i^{\text{th}}$  interval is

$$H(i) = \frac{j}{\sum_{k=i-j+1}^i S(k)}, \quad j < i < n. \quad (5)$$

This estimate will not work for  $i < j$ , but for the first  $j - 1$  intervals, one can let  $j = i$ .

#### Mixture Distributions Due to Parameter Drift

Reaction-time distributions, if they are to be useful for the present purpose, require a large number of trials which must be gathered in manageable blocks of trials over a period of many days. This extended data gathering raises doubts about the stability of

whatever parameters define the detection process (Grice, 1968). Parameter drift can have a pronounced effect on the tails of the reaction-time distributions, as may be shown by example. Suppose  $T_D$  has the simplest hazard function,  $h_D(t) = \lambda$ , so the distribution is the exponential  $\lambda e^{-\lambda t}$ . Suppose further that the parameter  $\lambda$ , rather than remaining constant, has some distribution, say, the gamma  $\mu^{n+1} \lambda^n e^{-\mu \lambda} / n!$ . It is not difficult to show<sup>2</sup> that the resulting mixture distribution has the hazard function  $h(t) = (n + 1) / (\mu + t)$ , which, unlike its component constant hazard functions, is a decreasing function of  $t$ . It would, therefore, exhibit the concave character we expect of a system composed of change and level detectors despite the fact that it is made up of nothing but a randomly responding detector with a drifting rate parameter.

Two types of parameter drift can be detected, and their effects can be largely negated by appropriate data-censoring methods. The first of these is a fore-period effect—systematic drift as a function of fore-period length. If there is no such drift, then, except for the impact of residual delays (e.g., muscular inertia and motoneuron conduction delays) on the first two or three hundred milliseconds, the false-alarm distributions should exhibit constant hazard rates. As will be seen later (Figure 4), the hazard rates begin at zero and approximate a steady value only after 500 to 750 msec for all three observers in all conditions. The rate is then constant up to about 9 sec, after which it increases for two observers and decreases for the third.

Parameters may also change from day to day or from one block of trials to the next. Evidence of such drift can be found in the variability of block mean reaction times (MRTs). If all parameters remained stable, the variability of block MRTs would necessarily obey familiar laws: Since reaction times are approximated by the lognormal distribution, the block geometric means should be distributed as a normal with a mean equal to the underlying geometric mean of the reaction-time distribution and a variance of  $V/N$ , where  $V$  is the underlying geometric variance of the reaction-time distribution and  $N$  is the number of trials in each block (excluding false alarms). But the observed variability in these block MRTs is much greater than expected if the parameters remain fixed across blocks.

A third type of drift, which is neither identifiable nor removable, is irregular parameter drift within a block of trials. The random criterion model to be discussed later (Grice, 1968, 1972) predicts a peaked hazard function precisely because it explicitly assumes that the reaction-time distribution is a mixture due to a randomly fluctuating criterion. Trial-by-trial feedback presumably aids the subject in stabilizing his criteria, but one can never rule out this sort of criterion shift simply by experimental manipulation.

### Data Censoring Methods

The aforementioned foreperiod effect can be dealt with simply: we merely reject from further analysis all trials on which the foreperiod was less than about 500 msec or more than 9,000 msec (the exact values depend on the observer). That is, we assume that the false-alarm rate is indicative of the subject's "readiness to respond." We reject trials for which the foreperiod was so short or so long that the subject may not have been appropriately ready to respond at the time the signal occurred.

The block-to-block drift presents a more difficult problem. We wish to eliminate as few blocks as possible in order to retain the maximum number of trials in the tails of the censored distributions, but we also must exclude blocks with extreme means in order to minimize the effect of mixtures on the tails. Therefore, we accept only those blocks with block means ( $M$ ) that satisfy the relation  $|Z_i| < c$ , where  $Z_i$  is the usual normal deviate  $|Z_i| = (M_i - M)/\sqrt{V_i/N_i}$ ,  $N_i$  is the number of RTs in the  $i$ th block, and  $c$  is a criterion. The larger the criterion, the more blocks are accepted, but the more serious the possible effect of mixture. We use values of  $c$  between 1.25 and 2, according to the need for enough data in the tails. (In 10 of 17 distributions,  $c$  was set at 1.5.) However, because the values of  $M$  and  $V$  are not known in advance, the censoring proceeds iteratively, beginning with the assumption that the correct values of  $M$  and  $V$  are just the mean and variance of the uncensored distribution. At each successive step, the true mean and variance of the censored distribution are assumed to be those resulting from the previous step. The process continues until the mean and variance obtained for the censored distribution closely approximate those assumed in the block selection process.

The use of geometric MRTs in the block-censoring poses one further problem because of the sensitivity of that statistic to very short reaction times. Responses that occur before the end of the foreperiod are clearly false alarms, as are responses that are initiated prior to the signal but completed after the signal. Moreover, if there is some nonzero minimum latency ( $t_m$ ) in the decision process, responses initiated before that minimum should also be treated as false alarms. The minimum observed response time with loud signals (which produce very few false alarms) was 140 msec, and, more importantly,  $t_m$  for the weak signals we used may add considerably to that minimum. Thus, if we knew the legitimate minimum response time, we could ignore all faster responses and thereby avoid excluding a block on the basis of an excessively small geometric MRT made so by happenstance. Here the hazard function is again useful. The legitimate minimum response time can be estimated as the point in the reaction-time distribution at which the hazard rate first rises above the false-alarm rate. The point cannot be determined with precision, but

it can be estimated to within 50 msec or so, and responses occurring prior to that point can be treated as false alarms. With the weak signals we used, the minimums were found to be between 200 and 250 msec, whereas for strong signals it is on the order of 150 msec. Assuming that the strong signal latency is due almost entirely to residual processes, the difference between the strong-signal minimum and the weak-signal minimum represents a minimum decision latency on the order of 50 to 100 msec. This is in agreement with a result of Green and Luce (1971).

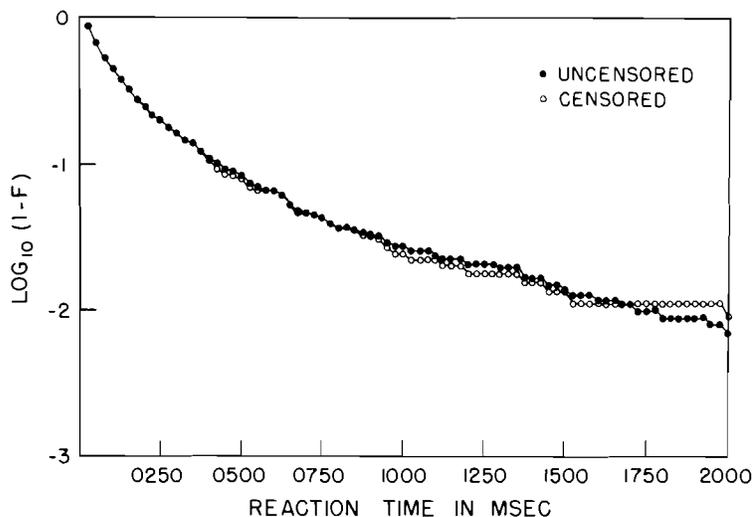
Block-to-block parameter drift, in addition to affecting the block MRT, also might be expected to affect the block false-alarm rate. We therefore took the precaution of excluding blocks with extreme false-alarm rates. The bounds were chosen by examining scatterplots of MRT vs. P(FA). These scatterplots typically showed a central cluster of blocks surrounded by sparsely scattered outliers. This censoring may well have been unnecessary, but if the variability in false-alarm rate was independent of any parameter drift, then exclusion of blocks according to false-alarm rate did no harm other than to decrease the number of trials in the final censored distributions. If, however, the variability is due in part to parameter drift, then censoring according to the false-alarm rate should reduce the effect of that drift.

### Evaluation of the Censoring Method

There are no established results to guide us in evaluating the censoring methods just described. It seemed advisable, therefore, to do a Monte Carlo study of such censoring on simulated data obtained from pure distributions (i.e., those with fixed parameters) and mixture distributions (where the parameters are varied deliberately). We chose to base the simulations on the Wald distribution (see Discussion section), which has a peaked hazard function.

The first test of the effects of censoring used 1,000 independent samples generated from a Wald distribution (arithmetic mean = 200, standard deviation = 350) and organized into 20 blocks of 50 trials each (a block size comparable to that of the reaction-time data). The criterion ( $c$ ) chosen for block exclusion was very conservative (.5 instead of the 1.5 which was typically used with the data) in order to provide a severe test of the censoring procedure. The results, shown in Figure 2, are rather surprising. Even though 55% of the blocks were rejected, the resulting sample log survivor function is practically unchanged.

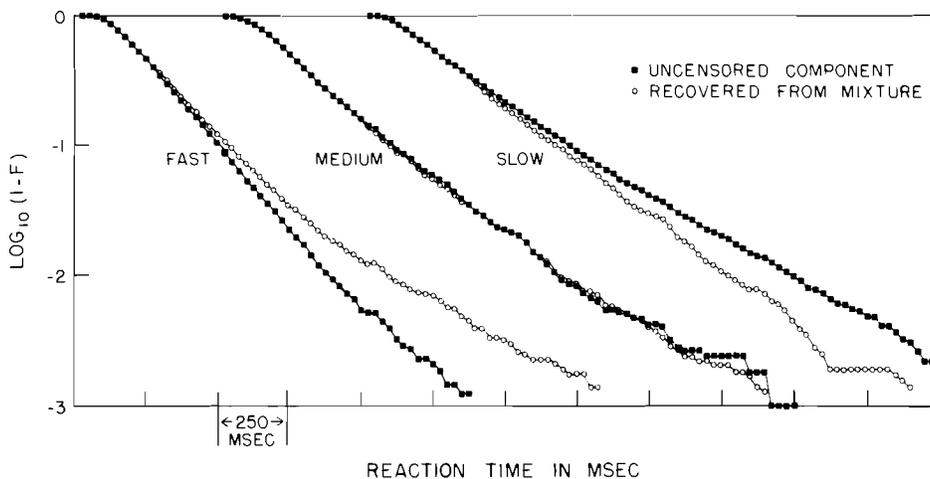
The second test concerned the effectiveness of the censoring process in extracting a reasonably pure distribution from a mixture made up of known distributions. To this end, we produced a distribution composed of a mixture of three known Wald distributions with arithmetic means and variances that are representative of values we obtained in fast, medium, and slow blocks of reaction-time data. The means



**Figure 2.** Effect of censoring on a known Wald distribution of true mean 200 and standard deviation 350. The sample mean was 189.88 and standard deviation 326.67. The block censoring criteria were: geometric mean ( $m$ ) = 89.76, standard deviation ( $s$ ) = 3.28, and criterion ( $c$ ) = .5.

were 280, 320, and 360 msec, and the standard deviations were, respectively, 180, 240, and 301. Ninety-nine blocks of 50 trials each were generated for each of the three distributions and pooled into one distribution. Our block-censoring process was then used to extract each of the three component distributions

(with the exception that we used the known means and variances rather than the iterative method). The results were not particularly sensitive to the value of the criterion ( $c$ ); two values, .50 and .84, were tried, and both gave similar results. The recovery of the medium distribution was quite good (Figure 3). How-



**Figure 3.** Effect of censoring mixed Wald distributions. The three components shown, with sample statistics shown below were combined into one mixed distribution. The summary statistics of the pure distributions are given at the top of each panel. The geometric mean and standard deviations of these pure distributions were used in succession in an attempt to recover each of the original components from the mixture (with  $c = .84$  in all three cases). The summary statistics are as follows:

	Uncensored		Censored	
	Mean	SD	Mean	SD
fast component	276.55	175.24	288.46	211.97
medium component	319.29	240.90	315.91	236.61
slow component	358.30	296.21	339.48	255.49
	273.59	2.06	269.43	1.97

ever, the process was unable to recover the fast or slow distribution, and the manner of failure is instructive. In both cases, the result of the censoring process tended to be biased toward the medium distribution. This bias is a valuable property of the censoring process if, as seems likely, the distribution of detection parameters over blocks is unimodal. Results of early iterations should be biased toward the unmixed distribution that would result from the modal parameters, and subsequent iterations would, if the data are as well behaved as our simulation, produce something very close to that modal reaction-time distribution.

## RESULTS AND DISCUSSION

### Hazard Functions of the Censored Distributions

Figure 4 shows examples of the log survivor function for false alarms, Figure 5 shows the estimated hazard functions of the censored distributions, and Table 1 gives the associated summary statistics.

Before embarking on a discussion of the form of the distributions in Figure 5, let us first consider the effect of frequency and intensity on the distributions. The intensity effect is consistent and surprisingly large, given that 2 dB constitutes at most only a few JNDs. At the higher intensity, the arithmetic MRT, averaged across subjects and frequencies, was 495 msec. At the lower intensity, it was 676 msec. The corresponding average standard deviation changed from 172 to 315 msec, nearly a factor of two. Most interesting, however, is the effect of intensity on the shape of the hazard functions. At the lower intensity, the hazard functions show either a relatively small peak or, in several cases, virtually no peak at all, which suggests that the change detector is near or below threshold. In every case, the addition of 2 dB to the signal produced a marked increase in the hazard rate between about 400 and 1,000 msec after onset. The increased intensity also tends to increase the hazard rate past 1,000 msec, but by a smaller factor. Indeed, in two cases (S.B. at 1,000 and 4,000 Hz), the asymptotic hazard rates are virtually the same at both intensities despite the appearance of a large peak between 400 and 600 msec with the higher intensity. If detection is due to the parallel operation of a change and a level detector, then the addition of 2 dB seems to benefit the change detector far more than the level detector.

The frequency effects are more subtle. There seems to be no major qualitative difference between 250 and 4,000 Hz. There are some quantitative differences in MRT, and there is some suggestion that the ratio of peak hazard rate to asymptotic hazard rate may be higher for the 4,000-Hz data than for the other two frequencies. But these quantitative differences may be due to intensity inasmuch as threshold is well known to be a function of frequency.

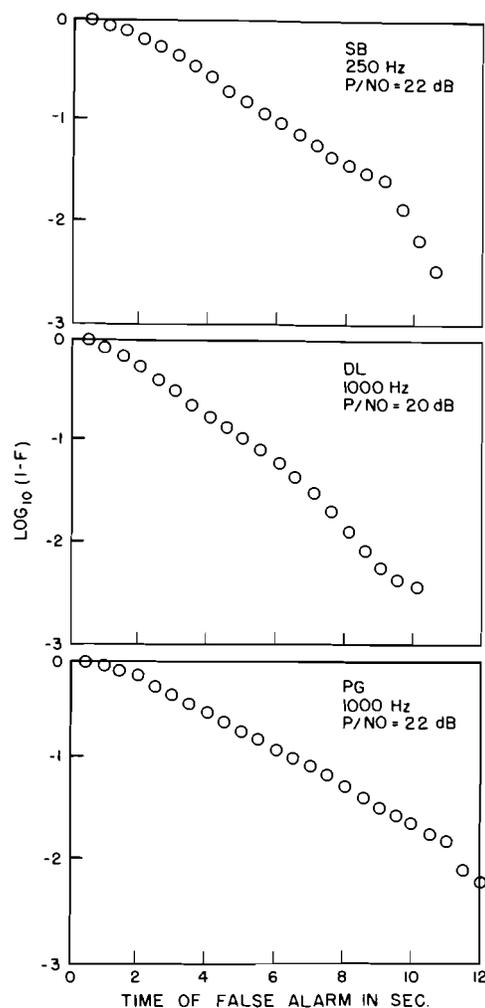


Figure 4. Examples of false-alarm distributions. Note that the false-alarm hazard rate is essentially zero for the first 500 msec.

The primary purpose of collecting these data was to determine the form of the reaction-time distributions. Some are clearly peaked, but it was not clear whether the tails of the hazard functions are constant, as would be predicted from the combination of both a change and a level detector. The Gnedenko test (Gnedenko, Belyayev, & Soloyev, 1969) with  $\alpha = .05$  indicates that the hazard functions of the tails are constant for all of the censored distributions but declining for six of the uncensored distributions (see Table 2). These declining rates are therefore presumably due to parameter drift. Thus, the conclusion is that the tails are asymptotically constant and substantially higher than zero.

### Other Models and Their Hazard Functions

A variety of distributions have been proposed for reaction time (see Table 3 for a summary). Some, for example, the log normal and the double monomial, are ad hoc attempts to account for the well-known

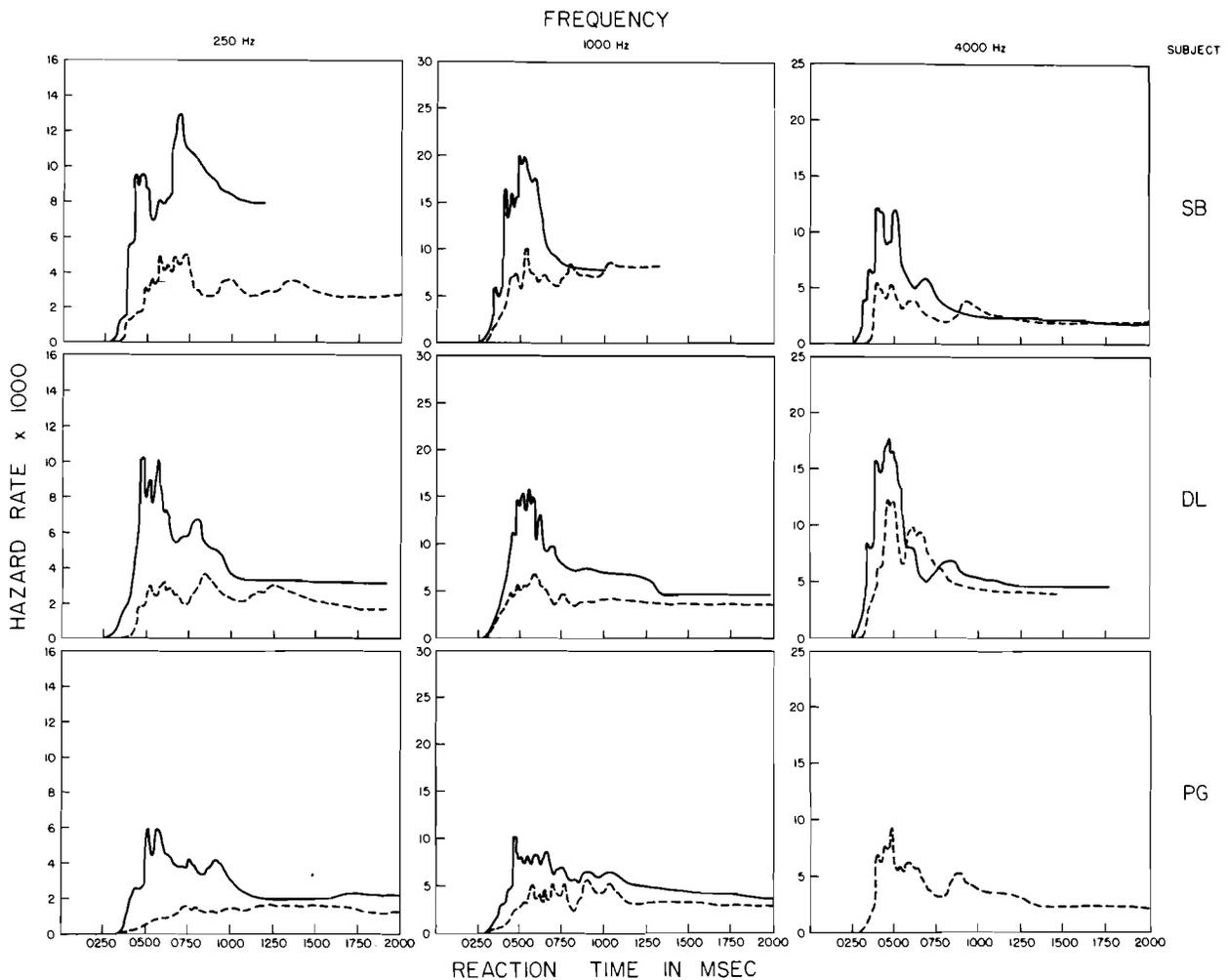


Figure 5. Estimated hazard functions of the censored distributions. The hazard functions were estimated by the random smoothing method with  $j = 20$ , and then the resulting curves were smoothed by eye so that two functions could be juxtaposed in each panel. The unsmoothed data are available in Burbeck (1979). The solid curve is the more intense signal, and the dotted one the less intense signal.

positive skew and high tails that characterize simple reaction-time distributions. Others, for example, the extreme value distributions and the one proposed by Thomas (1969), arise from models of choice reaction time and, although they may therefore not be relevant to simple reaction time, they are included for completeness.

Of those that have been derived from explicit models for simple reaction time, only two predict peaked hazard functions with asymptotically constant (non-zero) tails: (1) the Wald distribution (Wald, 1947) which arises from the random walk model (Emerson, 1970; McGill, 1963; Stone, 1960; Wald, 1947), and (2) Grice's random criterion model (1968, 1972). We therefore pay particular attention to these models. The Wald distribution is used in simulating some reaction-time data, and we attempt to fit it to the data as an alternative to change and level detectors.

Grice's model assumes that information about the stimulus accumulates deterministically—either linearly (Grice, 1968) or exponentially (Grice, 1972). Thus, for any given criterion, the reaction time is uniquely determined. In his model, the randomness found in reaction-time distributions is entirely due to a randomly fluctuating criterion which he assumes to be normally distributed. This model was proposed in the context of a fixed-foreperiod catch-trial paradigm and cannot account for the observed false-alarm behavior of subjects in a random-foreperiod paradigm: it predicts that all false alarms would occur at the instant the trial begins. Moreover, the assumption of deterministic information is not in accord with what we know about the peripheral auditory system. Nonetheless, the model highlights the potential problem of parameter drift and will be discussed further in a subsequent section.

**Table 1**  
**Summary Statistics of the Censored Reaction-Time Distributions**

Subject	N		Arithmetic		Geometric		P(FA)
	Original	Censored	Mean	SD	Mean	SD	
250 Hz, 20 dB							
S.B.	880	413	681.20	305.96	630.97	1.451	.16
D.L.	823	306	786.73	402.85	709.00	1.552	.26
P.G.	776	242	1108.71	608.39	959.05	1.725	.30
250 Hz, 22 dB							
S.B.	1629	417	481.20	120.16	467.44	1.265	.17
D.L.	1593	486	530.84	192.20	506.32	1.337	.19
P.G.	1426	385	673.35	341.54	618.24	1.464	.27
1,000 Hz, 20 dB							
S.B.	1852	514	500.77	151.58	480.62	1.327	.21
D.L.	1766	564	562.57	238.13	526.01	1.418	.31
P.G.	1688	700	674.29	273.81	629.41	1.437	.35
1,000 Hz, 22 dB							
S.B.	1848	593	414.75	86.04	406.88	1.212	.18
D.L.	2363	1041	474.82	122.25	463.52	1.234	.15
P.G.	2179	665	537.71	180.57	516.74	1.306	.23
4,000 Hz, 24 dB							
S.B.	948	240	629.55	385.78	557.21	1.576	.23
D.L.	970	269	462.00	155.89	442.46	1.324	.24
4,000 Hz, 26 dB							
S.B.	1490	635	444.17	230.88	416.43	1.367	.12
D.L.	2441	1353	403.15	105.58	393.02	1.239	.23
P.G.	952	332	530.42	285.80	489.67	1.430	.37
Noise							
S.B.	2034	627	167.80	12.33	167.30	1.075	.02
D.L.	2007	910	166.30	17.28	175.80	1.097	.02
P.G.	1977	595	199.60	28.21	198.00	1.130	.03

Note—There were no data for P.G. at 4,000 Hz, 24 dB.

### Fitting the Wald Distribution

Figure 6 shows that the Wald distribution is characterized by a peaked hazard function which becomes asymptotically constant. It should be noted, however, that the random walk model from which the Wald distribution is derived is neither a change nor a level detector in the sense that these are described in the introduction. As it is usually formulated for two-choice situations (Laming, 1968; Stone, 1960; Wald, 1947), the Wald sequential model begins at signal onset with numerical information about the signal (from a common random variable) being registered at discrete periods of time. When the sum of these numbers first falls outside an interval defined by two numerical barriers, one or the other response is initiated, corresponding to the barrier crossed. The tendency (expected value) of the information to cross one rather than the other barrier depends upon which signal is presented. In the Wald model, the random variable representing signal information is assumed to be Gaussian distributed. (It is not widely realized that the form of the distribution of first crossings of the barriers is quite sensitive to deviations from

this Gaussian assumption.) To adapt this model to simple reaction times, two things need to be done. First, since there is only one signal to be responded to, one barrier is placed at  $-\infty$ ; this assumption can be shown to lead to the so-called Wald distribution for crossings of the remaining barrier, provided that the increments in the random walk are small relative to the barrier location. Second, since signal onset or offset is the issue, the collection of information must begin at the warning signal, and not at the unknown time at which the reaction signal is presented. This means that during the foreperiod the random variable should have mean 0 and then at the reaction signal change to a positive mean. With this assumption, the process is more akin to a level detector than a change detector: the decision random variable (i.e., the sum of the signal random variables) estimates the product of the level of the signal and the time the signal has been present—thus it is a kind of accumulated energy detector. This model has been discussed qualitatively by Laming (1968, pp. 80-82), but to our knowledge it has not been investigated in detail mathematically. So we confine our attention to

Table 2  
Results of Gnedenko Test for Constant Hazard Rates  
in the Tails of the Distributions

Subject	Censored	Uncensored
250 Hz, 20 dB		
S.B.	Constant beyond 750*	Decreasing
D.L.	Constant beyond 1520	Decreasing
P.G.	Increasing out to 2005 Constant beyond 2005	Decreasing
250 Hz, 22 dB		
S.B.	Constant beyond 750*	Decreasing
D.L.	Constant beyond 750*	Decreasing
P.G.	Constant beyond 815	Decreasing
1,000 Hz, 20 dB		
S.B.	Constant beyond 750*	Constant beyond 750*
D.L.	Constant beyond 750*	Constant beyond 1340
P.G.	Constant beyond 1114	Constant beyond 1581
1,000 Hz, 22 dB		
S.B.	Constant beyond 750*	Constant beyond 903
D.L.	Constant beyond 866	Constant beyond 997
P.G.	Constant beyond 1030	Constant beyond 874
4,000 Hz, 24 dB		
S.B.	Constant beyond 836	Constant beyond 1572
D.L.	Constant beyond 785	Constant beyond 1653
4,000 Hz, 26 dB		
S.B.	Constant beyond 850	Constant beyond 934
D.L.	Constant beyond 750*	Constant beyond 1173
P.G.	Constant beyond 1077	Constant beyond 1062

Note—There were no data for P.G. at 4,000 Hz, 24 dB.  
\*Tests not made for RT < 750 msec.

the somewhat unrealistic one-barrier model with known signal offset, leading to the Wald distribution.

As has long been recognized (Donders, 1868/1969), the observed reaction times are not direct measures of the decision process alone: the detector latencies are always blurred by other residual random processes. Assuming that these effects are purely additive, as seems plausible, then

$$T_{RT} = T_D + T_R, \quad (6)$$

where  $T_{RT}$  is the observed reaction time,  $T_D$  the detector (or decision) latency, and  $T_R$  the residual latency.

We make two strong assumptions for which there is little evidence one way or the other. First, we assume  $T_D$  and  $T_R$  are independent random variables. Second, we assume that the distribution of  $T_R$  is independent of signal intensity.

By Equation 6 and the assumption that  $T_D$  and  $T_R$  are independent,

$$f_{RT}(t) = \int_0^t f_D(t-x)f_R(x) dx. \quad (7)$$

If the characteristic function is defined by

$$C(\omega) = E(e^{i\omega T}), \quad i = \sqrt{-1}, \quad (8)$$

then it is well known that

$$C_{RT}(\omega) = C_D(\omega)C_R(\omega). \quad (9)$$

The conversion of the convolution integral, Equation 7, into the multiplication of Equation 9 is a great convenience for studying reaction times (Christie & Luce, 1956; Green, 1971).

To estimate  $h_D$  somewhat more accurately than by  $h_{RT}$ , we need to know something about  $T_R$ . Here we invoke the second assumption, namely that the distribution of  $T_R$  is independent of signal intensity. It is well known that as intensity is increased, both the mean reaction time and its variance decrease markedly. For example, Green and Luce (1971) show that near threshold the ratio of the standard deviation to the mean is about 1, whereas at high intensities the ratio drops to 1/10th or less (Hopkins & Kristofferson, 1980, report standard deviations as small as 7 msec for mean latencies varying from 310 to 550 msec). Letting an asterisk denote these quantities at a high signal level, we have the approximation

$$T_{RT}^* \cong d + T_R, \quad (10)$$

where  $d$  is the mean detection time for a very intense signal. It is well known that Equation 10 implies

$$C_{RT}^*(\omega) \cong e^{id} C_R(\omega). \quad (11)$$

Substituting in Equation 9,

$$C_D(\omega) \cong C_{RT}(\omega)e^{id\omega}/C_{RT}^*(\omega). \quad (12)$$

We do not know  $d$ , but if we are willing to assume that it is very small, Equation 12 reduces to

$$C_D(\omega) \cong C_{RT}(\omega)/C_{RT}^*(\omega). \quad (13)$$

Inverting the transform yields an estimate of  $f_D$ . We use this deconvolution procedure to test the fit of a specific model of  $f_D$ . The deconvolution procedure may produce a somewhat distorted estimate of  $f_D$ , particularly if the independence assumptions are wrong. But the alternative is to explicitly model the entire reaction-time process from the eardrum to the fingertip—a formidable, if not hopeless, task.

Deconvolution has one unfortunate consequence: Although it produces useful estimates of the decision distribution, it does not preserve the information

Table 3  
Hazard Functions of Several Theoretical Reaction-Time Distributions

Name of Distribution	F(t)	h(t)	Growth of h(t)	Reaction-Time References
Exponential	$1 - e^{-\lambda t}, \lambda > 0, t \geq 0$	$\lambda$	constant	
Gaussian (Normal)	$\int_{-\infty}^t \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx$ $\sigma > 0$	-	increasing to $\infty$	
Double exponential (Extreme Value Type I)	$\exp\{-\exp[-\alpha(t-\beta)]\}$ $\alpha > 0$	$\frac{\alpha e^{-\alpha(t-\beta)} F(t)}{1 - F(t)}$	increasing to $\alpha$	Green & Luce (1973) Wandell & Luce (1978)
Logistic	$1/(1 + \exp[-(t-a)/\beta])$ $\beta > 0$	$\frac{1}{\beta} F(t)$	increasing	
Gamma (Pearson Type III)	$\int_{\gamma}^t \frac{(x-\gamma)^{\alpha-1} \exp[-(x-\gamma)/\beta]}{\beta^{\alpha} \Gamma(\alpha)} dx$ $t \geq \gamma, \alpha > 0, \beta > 0$	-	increasing for $\alpha > 1$ to $1/\beta$ decreasing for $\alpha < 1$ to $1/\beta$	Christie (1952) Bush & Mosteller (1955) Christie & Luce (1956) Luce (1960) Restle (1961) McGill (1963)
Weibull (Extreme value Type III)	$1 - \exp\left[-\left(\frac{t-t_0}{\alpha}\right)^c\right]$ $t \geq t_0, c > 0$	$\frac{c}{\alpha} \left(\frac{t-t_0}{\alpha}\right)^{c-1}$	Increasing for $c > 1$ to $\infty$ decreasing for $c < 1$ to 0	
Pareto	$1 - (k/t)^a, t \geq k > 0, a > 0$	$a/t$	decreasing to 0	Woodworth & Schlosberg (1954)
Log Gaussian	$\int_{\theta}^t \frac{d}{(x-\theta)\sqrt{2\pi}} \exp\left[-\frac{1}{2}(\gamma + \delta \log(x-\theta))^2\right] dx$ $t \geq \theta$	-	peaked decreasing to 0	
Wald (Inverse Gaussian)	$\left(\frac{\lambda}{2\pi t^3}\right)^{1/2} \exp\left[-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right]$ $t \geq 0, \lambda > 0, \mu > 0$	-	peaked decreasing to $\lambda/2\mu^2$	Stone (1960) Laming (1968) Link (1975)

Extreme Value Type II	$\exp[-(a/t)^k]$ $t > 0, a > 0, k > 0$	$\frac{F(t)}{1-F(t)} \frac{k a^k}{t^{k+1}}$	peaked decreasing to 0	Wandell & Luce (1978)
Double Monomial	$\frac{(\epsilon+1)(\delta-1)}{t_0(\delta+\epsilon)} \left\{ \left(\frac{t}{t_0}\right)^\delta, 0 \leq t \leq t_0 \right.$ $\left. \left(\frac{t}{t_0}\right)^{-\delta}, t_0 \leq t \right\}$	$\frac{\epsilon+1}{t_0} \left(\frac{t}{t_0}\right)^\epsilon$ $\frac{\delta+\epsilon}{\delta-1} \left(\frac{t}{t_0}\right)^{\delta+1}, 0 \leq t \leq t_0$ $\frac{\delta-1}{t}, t_0 < t$	peaked decreasing to 0	Snodgrass et al. (1967)
Ex-Gaussian	$\frac{\mu}{\sigma} e^{\mu a + \mu^2 \sigma^2 / 2} e^{-\mu t}$ $\times \left[ \Phi\left(\frac{t - \gamma - \mu \sigma^2}{\sigma}\right) - \Phi\left(-\frac{\gamma - \mu \sigma^2}{\sigma}\right) \right]$ where $\Phi$ is unit Gaussian	-	increasing	Hohle (1965) Ratcliff & Murdock (1976) Ratcliff (1978)
(Unnamed)	$\Phi(a - b e^{-\beta t})$ , where $\Phi$ is unit Gaussian	-	peaked decreasing to $\beta$	Grice (1972)
Pearson Type V	$\int_{1/t}^{\infty} \beta^k x^{k-1} e^{-\beta x} \frac{dx}{\Gamma(k)}$	-	decreasing to 0	Thomas (1969)

Note—Although  $h(t)$  is often difficult to compute explicitly, its qualitative behavior can often be found by computing  $f'(t)/f(t) = (d/dt) \log f(t)$ . The asymptote is computed from

$$\lim_{t \rightarrow 0} h(t) = \frac{-\lim_{t \rightarrow 0} f'(t)}{t \rightarrow 0 f(t)}$$

The direction of  $h(t)$  is found using the following theorem of Thomas (1971). Suppose  $f$  is differentiable and  $f(t) \rightarrow 0$  as  $F(t) \rightarrow 1$ . If  $-\ln f(t)$  is convex (concave), then  $h(t)$  is increasing (decreasing).

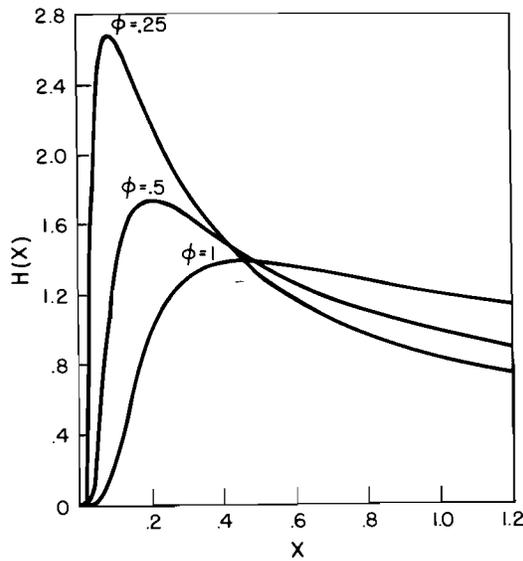


Figure 6. The hazard function of the Wald distribution. Each distribution has unit mean. The shape parameter,  $\phi$ , determines the sharpness of the peak in the hazard function. As  $\phi$  becomes large, the Wald distribution approaches the normal distribution, which has a monotonic increasing hazard function.

about individual reaction times that is required by the random smoothing method of hazard function estimation. The slopes of the deconvolved log-survivor functions indicate that deconvolution does not much change the shape of the hazard functions; it narrows the peaks somewhat, but primarily it results in a translation of the function toward the origin. Thus, because other methods of estimating the hazard functions are not very satisfactory, we present only the deconvolved log-survivor functions and leave the reader to verify that the deconvolved hazard functions retain their general shape.

With the present data, the individual reaction times were collapsed into 2-msec bins before transforming the histograms. The transforms of the weak-signal data were then divided by the transform of the loud-signal data to obtain the transforms of the estimated decision distributions. This process produces high-frequency components that reflect, primarily, the discontinuities inherent in the histograms. So, to reduce this problem and to arrive at reasonably smooth estimates of the decision distribution, all frequencies greater than 20 Hz were removed (for a discussion of this problem, see Davis, 1975, 1977). The filtered transforms were then inverted to obtain estimates of the decision histograms.

**Fitting the Wald Distribution**

The general form of the Wald distribution is

$$f(t) = [\mu\phi/(2\pi t^3)]^{1/2} \cdot \exp[\phi - (1/2)\phi(t/\mu + \mu/t)], \quad (14)$$

where  $\phi = \lambda/\mu$ , and  $\text{var}(X) = \mu^3/\lambda$ .

The maximum likelihood estimators for  $\mu$  and  $\lambda$  have been determined by Tweedie (1957), but they require raw data (i.e., individual reaction times) which are lost in the process of deconvolution. Moreover, the censored distributions are doubly truncated prior to deconvolution—reaction times longer than 3,000 msec or shorter than the estimated minimum legitimate reaction time were ignored—and the negative loops introduced in the tails by the deconvolution required further truncation. It is therefore convenient that Johnson and Kotz (1970, Vol. 1, p. 146) give estimators (due to Patel, 1965) that begin with sample moments from the doubly truncated distribution. Using these, we can, under the assumption that the data are from a Wald distribution, estimate  $\mu$  and  $\lambda$  and then determine the goodness of fit between the resulting Wald distribution and the data. Table 4 gives the results of that test.

The results of chi-square tests indicate that the Wald distribution, as estimated by Patel's method, does not fit the data. But, because slight alterations in the estimated parameters may produce smaller chi-square values, a trial and error gradient search program was used to find the minimum chi-square fit, beginning at the estimated parameters found by Patel's method. The minimum chi-square fits are somewhat better (see Table 4), but are still not satisfactory. Figure 7 shows two examples of deconvolved data together with the Wald distributions estimated by each method.

Although the Wald distribution provides a reasonably good account of the 1,000-Hz, 20-dB data of all subjects, it fails badly for the other five conditions. The source of the failure is the inability of the Wald distribution to account for the decreasing hazard rates in the tails. That may seem paradoxical inasmuch as the hazard functions of Wald distributions

Table 4  
Goodness of Fit Between the Deconvolved Distributions and Wald Distributions

Condition		Patel's Method		Minimum $\chi^2$	
Hz	dB	$\chi^2$	df	$\chi^2$	df
S.B.					
250	20	53.8	24	45.9	25
250	22	159.3	22	147.2	23
1000	20	29.4	18	26.6	17
1000	22	43.2	9	31.3	9
4000	24	1523.4	14	75.1	20
4000	26	127.3	11	85.2	11
P.G.					
250	20	*		143.7	31
250	22	129.9	19	125.8	19
1000	20	129.8	88	116.3	84
1000	22	96.8	31	81.9	32
4000	26	119.3	6	340.1	24

\*Cannot estimate because matrix is singular.

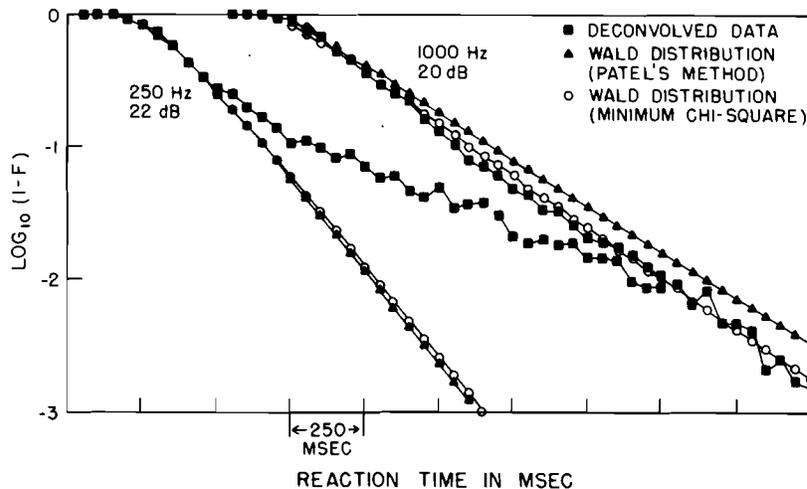


Figure 7. Samples of deconvolved data for PG vs. estimated Wald distributions. Note: Because of the negative excursions in the tails of the deconvolved histograms, the log survivor functions are not monotone decreasing.

are peaked. But the shape parameter,  $\phi$ , which determines the sharpness of the peaks, is linked to the coefficient of variation (cv), that is,  $(cv)^2 = 1/\phi$ . Figure 6 shows that the hazard function of the Wald distribution is markedly peaked only for  $\phi < 1$ , that is, for coefficients of variation larger than one. Moreover, the hazard function of the Wald distribution becomes more peaked as  $\phi$  decreases (or the coefficient of variation increases). The data, on the other hand, show coefficients of variation less than one, and smaller coefficients correspond to more peaked hazard functions. Therefore, the inability of the Wald distribution to account for the data is related to a fundamental difference between its properties and those of the data. But the failure of the Wald distributions to account for our data does not allow us to rule out the random walk model entirely, for the Wald is only the asymptotic form of the model. If the random walk is assumed to proceed in relatively few, large jumps, then it does not give rise to the Wald distribution. Perhaps, in that case, the predicted decision distribution has better properties, but it is not the purpose of the present paper to pursue the matter further.

#### Random Criterion Models

A random criterion model of the sort proposed by Grice may yet be able to account for our data. Although the model, as he formulated it, cannot account for random foreperiod data, it could be modified in a number of ways that would work. The resulting predicted hazard functions would not be exactly the same, but they would very likely be of the same general character, because mixtures (as produced by a random criterion) tend to give rise to peaked hazard functions. The model as given in Ta-

ble 1 has three free parameters, and modifications necessary to introduce a random component to the information accumulation would no doubt introduce more free parameters. Thus, it seems safe to assume that a random criterion model can be made to fit the data. However, the fact that the addition of 2 dB can change the hazard function from monotone increasing to strongly peaked (e.g., with Subject P.G. at 250 Hz) casts some doubt on the random criterion model. Mixtures of monotone increasing hazard functions can become peaked, but mixtures of peaked hazard functions cannot become monotone increasing (cf. Barlow & Proschan, 1965). Thus, the presence of both implies that the distribution for any fixed criterion is monotone increasing, and therefore that the height of the peak is a measure of the variance of the criterion. The data imply, then, that the criterion is more stable with a 20-dB signal than with a 22-dB signal—a curious result with little intuitive appeal.

Grice (1972) also presented reaction-time data which his model fits reasonably well. However, his data do diverge from his model, and the divergence suggests that the data has more strongly peaked hazard functions than the model. And, to the extent that such a fit implies that his data would show peaked hazard functions, a combined change and level detector model may account for his data just as well. Moreover, other sources of psychophysical evidence for change detection (referenced in the introduction) are not explainable by random criteria. Thus, it is not at all clear that a random criterion model is necessary to account for reaction-time data. Nor is it desirable. Some variability in criteria is no doubt inevitable in real world situations. But, if substantial variability is inescapable, we cannot hope to uncover much,

if anything, else about the decision process: it will always be confounded with whatever process controls the criterion.

### SUMMARY

Our analysis of reaction times to response-terminated, weak-signal offsets has rested upon two major methodological procedures. First, there are the experimental and analytical procedures designed to ensure that the distributions are built up from component distributions having minimal mixture artifacts. The data-censoring procedures are somewhat suspect, but we believe that the arguments presented and the simulations performed justify them. The second is the use of Fourier transform techniques and reaction-time distributions with loud signals to factor out the obscuring residual latencies. Although it is no doubt true that the loud-signal data include some decision component, those distributions are narrow enough for small changes in their shapes not to appreciably change the resulting estimates of the decision distributions. Thus, if the assumptions underlying the use of deconvolution are not badly wrong, then our estimates of the decision distributions ought to be reasonably good.

We find that the decision latency distributions show two forms, either monotone increasing hazard functions or peaked hazard functions with asymptotically constant tails. Moreover, these two forms are systematically related to the intensity of the signal: monotone increasing hazard functions are found only with the weaker signals. Although the data do not rule out a random criterion model, they are most consistent with the hypothesis that two types of detectors function independently, with detection being initiated by whichever detector responds first to the offset. The level detector responds somewhat sluggishly but continues to provide information after the offset and accounts for the asymptotic hazard rate. The change detector is quicker, is inherently less persistent, has a somewhat higher threshold, and accounts for the peaks of the hazard functions. Thus, the reaction-time data are in accord with the accumulating physiological evidence that the parallel operation of change and level detectors is a frequently encountered principle of neural organization.

### REFERENCE NOTE

1. Shannon, R. V. Unpublished data, 1973. Examples of these data appear in Burbeck (1979).

### REFERENCES

- AHUMADA, A., JR., MARKEN, R., & SANDUSKY, A. Time and frequency analysis of auditory signal detection. *Journal of the Acoustical Society of America*, 1975, **57**, 385-390.
- BARLOW, R. E., & PROSCHAN, F. *Mathematical theory of reliability*. New York: Wiley, 1965.
- BURBECK, S. L. *Change and level detectors inferred from simple reaction time*. Doctoral dissertation, University of California at Irvine, 1979.
- BUSH, R. R., & MOSTELLER, F. *Stochastic models for learning*. New York: Wiley, 1955.
- CARTERETTE, E. C., FRIEDMAN, M. P., & COSMIDES, R. Reaction-time distributions in the detection of weak signals in noise. *Journal of the Acoustical Society of America*, 1965, **38**, 531-542.
- CHRISTIE, L. S. The measurement of discriminative behavior. *Psychological Review*, 1952, **59**, 443-452.
- CHRISTIE, L. S., & LUCE, R. D. Decision structure and time relations in simple choice behavior. *Bulletin of Mathematical Biophysics*, 1956, **18**, 89-112.
- CLELAND, B. G., DUBIN, M. W., & LEVICK, W. R. Sustained and transient cells in the cat's retina and lateral geniculate nucleus. *Journal of Physiology (London)*, 1971, **217**, 473-496.
- DAVIS, K. B. Mean square error properties of density estimates. *The Annals of Statistics*, 1975, **3**, 1025-1030.
- DAVIS, K. B. Mean integrated square error properties of density estimates. *The Annals of Statistics*, 1977, **5**, 530-535.
- DONDERS, F. C. [On the speed of mental processes.] In W. G. Koster (Ed. and trans.), *Attention and performance II*. Amsterdam: North-Holland, 1969. (Reprinted from *Acta Psychologica*, 1969, **30**.) (Originally published 1868.)
- EMERSON, P. L. Simple reaction time with markovian evolution of gaussian discriminial processes. *Psychometrika*, 1970, **35**, 99-109.
- ENROTH-CUGELL, C., & ROBSON, J. G. The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology (London)*, 1966, **187**, 517-552.
- GERSTEIN, G. L., BUTLER, R. A., & ERULKAR, S. D. Excitation and inhibition in cochlear nucleus. I. Tone-burst stimulation. *Journal of Neurophysiology*, 1968, **31**, 526-536.
- GNEDENKO, B. V., BELYAYEV, YU. K., & SOLOYEV, A. D. *Mathematical methods of reliability theory*. (Translated by Scripta Technica; translation edited by Richard E. Barlow.) New York: Academic Press, 1969.
- GREEN, D. M. Fourier analysis of reaction time data. *Behavior Research Methods & Instruments*, 1971, **3**, 121-125.
- GREEN, D. M., & LUCE, R. D. Detection of auditory signals presented at random times: III. *Perception & 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 IV*. New York: Academic Press, 1973.
- GRICE, G. R. Stimulus intensity and response evocation. *Psychological Review*, 1968, **75**, 359-373.
- GRICE, G. R. Application of a variable criterion model to auditory reaction time as a function of the type of catch trial. *Perception & Psychophysics*, 1972, **12**, 103-107.
- HOHLE, R. H. Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, 1965, **69**, 382-386.
- HOPKINS, G. W., & KRISTOFFERSON, A. B. Ultrastable stimulus-response latencies: Acquisition and stimulus control. *Perception & Psychophysics*, 1980, **27**, 241-250.
- JOHNSON, N. L., & KOTZ, S. *Continuous univariate distributions* (2 vols.). Boston: Houghton Mifflin, 1970.
- KIANG, N. Y. S., WATANABE, T., THOMAS, E. C., & CLARK, L. F. *Discharge patterns of single fibers in the cat's auditory nerve*. Cambridge, Mass: M.I.T. Press, 1965.
- LAMING, D. R. J. *Information theory of choice reaction times*. London and New York: Academic Press, 1968.
- LINK, S. W. The relative judgment theory of two choice reaction times. *Journal of Mathematical Psychology*, 1975, **12**, 114-135.
- LUCE, R. D. Response latencies and probabilities. In K. A. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences*, 1959. Stanford: Stanford University Press, 1960.
- LUCE, R. D., & GREEN, D. M. Detection of auditory signals presented at random times, II. *Perception & Psychophysics*, 1970, **7**, 1-14.

- LUCE, R. D., & GREEN, D. M. A neural timing theory for response times and the psychophysics of intensity. *Psychological Review*, 1972, 79, 14-57.
- MACMILLAN, N. A. Detection and recognition of increments and decrements in auditory intensity. *Perception & Psychophysics*, 1971, 10, 233-238.
- MACMILLAN, N. A. Detection and recognition of intensity changes in tone and noise: The detection-recognition disparity. *Perception & Psychophysics*, 1973, 13, 65-75.
- MARROCCO, R. T. Sustained and transient cells in monkey lateral geniculate nucleus: Conduction velocities and response properties. *Journal of Neurophysiology*, 1976, 39, 340-353.
- MCGILL, W. J. Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*. New York: Wiley, 1963.
- MCGILL, W. J., & GIBBON, J. The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, 1965, 2, 1-18.
- MILLER, D. R., & SINGPURWALLA, N. D. *Failure rate estimation using random smoothing*. National Technical Information Service No. AD-A040999/SST, 1977.
- MILLER, G. A. The perception of short bursts of noise. *Journal of the Acoustical Society of America*, 1948, 20, 160-170.
- MØLLER, A. R. Unit responses in the cochlear nucleus of the rat to pure tones. *Acta Psychologica Scandinavica*, 1969, 75, 530-541.
- NICKERSON, R. S. Expectancy, waiting time and the psychological refractory period. *Acta Psychologica*, 1967, 27, 23-34.
- PATEL, R. C. Estimates of parameters of truncated inverse Gaussian distribution. *Annals of the Institute of Statistical Mathematics*, Tokyo, 1965, 17, 29-33.
- PFEIFFER, R. R. Classification of response patterns of spike discharges for units in the cochlear nucleus: Tone burst stimulation. *Experimental Brain Research*, 1966, 1, 220-235.
- RATCLIFF, R. A theory of memory retrieval. *Psychological Review*, 1978, 85, 59-108.
- RATCLIFF, R., & MURDOCK, B. B., JR. Retrieval processes in recognition memory. *Psychological Review*, 1976, 83, 190-214.
- RESTLE, F. *Psychology of judgment and choice*. New York: Wiley, 1961.
- RICE, J., & ROSENBLATT, M. Estimators of the log survivor function and hazard function. *Sankhya: The Indian Journal of Statistics*, Series A, 1976, 38, 60-78.
- 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.
- SMALL, A. M. Periodicity pitch. In J. V. Tobias (Ed.), *Foundations of modern auditory theory* (Vol. 1). New York: Academic Press, 1970.
- SNODGRASS, J. G. Foreperiod effects in simple reaction time: Anticipation or expectancy? *Journal of Experimental Psychology Monograph*, 1969, 79, 1-19.
- SNODGRASS, J. G., LUCE, R. D., & GALANTER, E. Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, 1967, 75, 1-17.
- STONE, M. Models for choice-reaction time. *Psychometrika*, 1960, 25, 251-260.
- THOMAS, E. A. C. Alternative models for information processing: Constructing non-parametric tests. *British Journal of Mathematical and Statistical Psychology*, 1969, 22, Part 2, 105-113.
- THOMAS, E. A. C. Sufficient conditions for monotone hazard rate: An application to latency-probability curves. *Journal of Mathematical Psychology*, 1971, 8, 33-332.
- TOLHURST, D. J. Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Research*, 1975, 15, 1143-1149.
- TWEEDIE, M. C. K. Statistical properties of inverse Gaussian distribution, I. *Annals of Mathematical Statistics*, 1957, 28, 362-377.
- WALD, A. *Sequential analysis*. New York: Wiley, 1947.
- WANDELL, B., & LUCE, R. D. Pooling peripheral information: Average versus extreme values. *Journal of Mathematical Psychology*, 1978, 17, 220-235.
- WATSON, G. S., & LEADBETTER, M. R. Hazard analysis I. *Biometrika*, 1964, 51, 175-184. (a)
- WATSON, G. S., & LEADBETTER, M. R. Hazard analysis II. *Sankhya*, Series A, 1964, 26, 101-116. (b)
- WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology* (rev. ed.). New York: Holt, 1954.

## NOTES

1. By independence,  $1 - F(t) = [1 - F_c(t)][1 - F_f(t)]$ . Taking the derivative,  $f(t) = f_c(t)[1 - F_f(t)] + f_f(t)[1 - F_c(t)]$ . The result is immediate.

2.

$$\begin{aligned} f(t) &= \int_0^{\infty} \lambda e^{-\lambda t} \mu^{n+1} \lambda^n e^{-\mu \lambda} d\lambda / n! \\ &= (\mu^{n+1} / n!) \int_0^{\infty} \lambda^{n+1} e^{-\lambda(\mu+t)} d\lambda \\ &= (n+1) \mu^{n+1} / (\mu+t)^{n+2} \end{aligned}$$

integrating to get  $1 - F(t)$  and solving for  $h(t)$  yields the result.

3. The geometric standard deviation is simply the antilog of the arithmetic sample standard deviation of the logs of the reaction times.

(Manuscript received December 1, 1980;  
revision accepted for publication March 24, 1982.)