

# Testing the assumptions of exponential, additive reaction time models

F. GREGORY ASHBY

Ohio State University, Columbus, Ohio 43210

Two assumptions commonly made by choice reaction time (RT) models are (1) that certain experimental tasks can be found that cause an extra processing stage to be inserted into the cognitive process and (2) that the duration of one or more processing stages is random with an exponential distribution. Few rigorous tests of either assumption exist. This paper reviews existing tests and presents several new results that can be used to test these assumptions. First, in the case in which the duration of an inserted stage is exponentially distributed, it is shown that the observable RT density functions must always intersect at the mode of the density requiring the extra processing stage. Second, when only the first assumption (Assumption 1) is made, it is shown that the cumulative RT distribution functions and, in many cases, the hazard functions must be ordered. Finally, when only Assumption 2 is assumed, it is shown that, under fairly weak conditions, the tail of the RT density function must be exponential. The first two results are applied to data from a memory scanning experiment, and the two assumptions are found to receive tentative support.

An assumption found in many early theories of choice reaction time (RT) is that two experimental tasks can be found such that the second task requires all the cognitive processes or stages of the first, plus one additional stage. It is also usually required that the inserted stage have no effect on the duration of any other stage. This assumption, popularized by Donders (1969) in 1868, has been termed the assumption of pure insertion, and it formed the basis of Donders' method of subtraction. His idea was to estimate the duration of the inserted stage by subtracting the mean RTs of the two tasks from each other. After a brief period of popularity, the method of subtraction fell into disfavor; its downfall proved to be introspectionistic reports, coming first from Wundt's laboratory and later from Külpe's in Würzburg, that the tasks recommended by Donders evoked qualitatively different experiences. Trained observers did not report experiencing concatenated mental processes, as pure insertion predicts.

Despite the subsequent loss of confidence in introspection, the bias against the subtractive method remained. A telling example of this attitude can be found in Boring's (1950) influential *History of Experi-*

*mental Psychology*, in which the method is dismissed on what are essentially these same subjective grounds. One seemingly more objective criticism of the assumption is that, at times, there is a large variability in the obtained estimates of the inserted stage (mean) duration (e.g., Boring, 1963; Smith, 1968). While this finding may cause a loss of confidence in subtractive method estimates, it has nothing to do with the validity of the pure insertion assumption. If there is a large variability in the duration of the inserted stage, there will be a large variability in the subtractive method estimate, regardless of whether pure insertion holds.

Despite these disclaimers, in certain experimental paradigms the assumption of pure insertion is still heavily relied upon. Examples include signal detection (e.g., Luce & Green, 1972; Kohfeld, Santee, & Wallace, Note 1), memory scanning (e.g., Sternberg, 1966), visual search (e.g., Atkinson, Holmgren, & Juola, 1969), and the Posner and Mitchell (1967) letter identification task (e.g., Hunt, 1978). Other examples include developmental studies in which subtraction is used to correct for differences in response execution time across different age populations (e.g., Surwillo, 1977). Thus, in spite of its many critics, the assumption of pure insertion enjoys continued popularity. A more objective test of the pure insertion hypothesis is a major goal of this paper.

A second assumption popular in RT theories is that the duration of one or more of the processing stages is random with an exponential distribution (e.g., Christie & Luce, 1956; Green & Luce, 1971; Hohle, 1965; Luce & Green, 1970; McGill, 1963; McGill & Gibbon, 1965; Restle, 1961; Taylor, 1976; Townsend, 1972, 1974, 1976). This seems to be the most commonly

Some of these results were presented at the Thirteenth Annual Mathematical Psychology Meetings in Madison, Wisconsin, August 1980, and some were included in a doctoral dissertation submitted to Purdue University. I would like to thank the members of my graduate advisory committee, James T. Townsend, Peter H. Schönemann, Richard Schweickert, Robert D. Sorkin, and Ben J. Winer, as well as Steven W. Link, R. Duncan Luce, and David Noreen for their helpful comments. Requests for reprints should be sent to F. Gregory Ashby, Human Performance Center, Ohio State University, 404B West 17th Avenue, Columbus, Ohio 43210.

made distributional assumption about the duration of RT components. Much of its popularity is undoubtedly due to its mathematical simplicity, but despite its prevalence, it too has been attacked on what are largely subjective grounds. Sternberg (1975) criticized the assumption because the exponential distribution is used as a model of waiting times and thus, in his words, "is fundamentally inconsistent with what we normally mean by the concept of processing over time" (p. 18).

The influence of these subjective criticisms can be attributed to the lack of adequate tests in the case of both assumptions. By far the oldest and best known test of pure insertion is to examine the difference between observed mean RTs from the two experimental conditions. If the mean from the condition thought to require the extra processing stage is greater than the mean from the other condition, then pure insertion is supported. Only if the means are ordered in the opposite direction can the assumption be rejected. Unfortunately, this is a weak test, since the experimental task thought to require the extra processing stage is invariably the more complex of the two.

A more rigorous approach is to first assume pure insertion holds and then try to estimate the probability distribution of the inserted stage duration. If the resulting estimate is reasonable, then pure insertion is supported; if it is not, the assumption can be rejected. As will be seen below, there are several variants of this potentially very powerful approach (e.g., Green, 1971; Green & Luce, 1971; Kohfeld et al., Note 1; Sternberg, Note 2), but unfortunately, they tend to be computationally complex and to be associated with some rather thorny statistical problems that are not yet completely resolved.

The assumption that at least some RT components are exponentially distributed is in much the same straits. No completely acceptable test exists. One possibility is based on the fact that if all RT components are exponentially distributed or if all RT components not exponentially distributed have a maximum time, then the tail of the observable RT density function is exponential (Green & Luce, 1971; Luce & Green, 1970; McGill & Gibbon, 1965). Indeed, considerable evidence exists that, at least in signal detection tasks in which weak signals are used, the tails of empirical RT densities are exponential (e.g., Green & Luce, 1967, 1971; Luce & Green, 1970; McGill, 1963; McGill & Gibbon, 1965; Burbeck & Luce, Note 3). Unfortunately, the above conditions under which the tail is known to be exponential are satisfied by a fairly small set of choice RT models. These conditions will be weakened below, where it will be shown that a much larger class of models predict such tails.

A second approach has been to fit a theoretical RT distribution containing an exponential component to an empirically obtained distribution (Hohle, 1965; Ratcliff, 1978; Ratcliff & Murdock, 1976; Snodgrass, Luce, & Galanter, 1967). A good fit supports the exponential

assumption. Hohle (1965) had good success fitting the convolution of a normal and an exponential distribution to data from a signal detection task, and Ratcliff (1978; Ratcliff & Murdock, 1976) found this same family of distributions gave good fits to empirical RT distributions obtained in memory experiments. One drawback of this procedure, of course, is that it requires precise specification of all RT components and not just the exponential one.

Recently, Ashby and Townsend (1980) derived a testable condition that is predicted by all models assuming both pure insertion and that the duration of the inserted stage is exponentially distributed. The proposed test is very strong, since, in addition, the condition can be predicted only by models making these assumptions. While the test is not particularly difficult to implement, neither is it trivial to do so. The next section of this paper details a second test of these two assumptions that is not quite as strong as the Ashby and Townsend (1980) test but has the advantage of being extremely easy to apply. It is therefore meant to play a diagnostic role, allowing one to quickly reject exponential, additive models when they are not appropriate. When the data do not allow this class of models to be rejected, an appeal to the stronger Ashby and Townsend (1980) test is recommended, since its resolving powers are greater.

In the third section of this paper, the consequences of dropping the exponential assumption will be examined and some alternative tests of models assuming only pure insertion will be suggested. Finally, the last section briefly considers tests of the exponential assumption that do not require pure insertion to hold.

#### A TRANSMODAL PROPERTY OF EXPONENTIAL, ADDITIVE LATENCY MODELS

Let me begin by formalizing the assumptions a bit. First, denote the observable RT in the experimental condition thought to require the extra processing stage as  $RT_k$  and the observable RT in the simpler task as  $RT_{k-1}$ . Finally, call the duration of the unobservable inserted stage  $T_k$ . While the additivities predicted by pure insertion might occur only at the level of mean RTs, I will assume that pure insertion imposes additivities at the distributional level. Thus the assumption of pure insertion at the distributional level can be stated for  $k > 1$  as

$$RT_k = RT_{k-1} + T_k, \quad (1)$$

where  $RT_{k-1}$  and  $T_k$  are independent random variables.<sup>1</sup> The idea of the subscripts here is that at some point we may wish to consider the possibility of a third experimental condition, requiring one less processing stage than Condition  $k - 1$ . With the present notation, the observable RT in this new condition can be called  $RT_{k-2}$ .

Equation 1, together with the independence assump-

tion, implies additivity, not only of the means, but also of all cumulants (e.g., Parzen, 1960). In addition, it implies that the observable RT density functions,  $g_k(t)$  and  $g_{k-1}(t)$ , from the two experimental conditions are related by

$$g_k(t) = \int_0^t g_{k-1}(t - t_0) f_k(t_0) dt_0 = g_{k-1}(t) * f_k(t), \quad (2)$$

where  $f_k(t)$  is the density function of the unobservable inserted stage duration. Thus, pure insertion at the distributional level implies that the observable RT density,  $g_k(t)$ , is the convolution of the observable density  $g_{k-1}(t)$  and the unobservable  $f_k(t)$ .

The assumption that the inserted stage duration is exponentially distributed with rate  $V_k$  is formalized as

$$f_k(t) = V_k \exp(-V_k t). \quad (3)$$

This exponential assumption imposes no restrictions on the distribution of  $RT_{k-1}$ , since it applies only to  $T_k$ , the duration of the inserted stage. In general, the results to be described in this paper will hold for any well-defined density function,  $g_{k-1}(t)$ .

Under these two assumptions, Ashby and Townsend (1980) showed that

$$V_k = \frac{g_k(t)}{G_{k-1}(t) - G_k(t)}, \quad \text{for all } t > 0, \quad (4)$$

where  $G_k(t)$  is the cumulative distribution function associated with  $g_k(t)$ . A test of the assumptions is possible, since the left-hand side of Equation 4 is time invariant and thus a plot of  $g_k(t)/[G_{k-1}(t) - G_k(t)]$  vs. time should be constant if Equations 2 and 3 are true. Further, it turns out that if such a plot is constant, the two assumptions must be satisfied by the data from the two experimental conditions. Ashby and Townsend (1980) also suggest statistical decision rules that can be utilized when implementing such a test.

As it happens, there is another very simple way of telling if the assumptions embodied in Equations 2 and 3 are violated, which merely involves plotting the empirical RT density function estimates and checking their points of intersection.

**Theorem 1:** If pure insertion holds at the distributional level and if the duration of the inserted stage is exponentially distributed, then  $g_k(t)$  and  $g_{k-1}(t)$  intersect at the maxima and minima of  $g_k(t)$  and only at these points.

*Proof:* The proofs of all results in this paper are contained in the appendix.

When  $g_k(t)$  is unimodal, there is, of course, only one maximum (in conventional cases, consideration of minima can be excluded), and so the theorem requires the two densities,  $g_k$  and  $g_{k-1}$ , to intersect at the

mode of  $g_k(t)$  and only there. Prior to that mode,  $g_k(t) < g_{k-1}(t)$ , whereas after the mode this ordering is reversed. This situation is illustrated in Figure 1, in which both  $g_k$  and  $g_{k-1}$  are gamma densities with the same rate. The leftmost curve,  $g_{k-1}(t)$ , is a two-stage gamma, and the rightmost curve,  $g_k(t)$ , has three stages (thus,  $k = 3$ ). Clearly, the assumptions of Theorem 1 are met in this case, and the intersection of the two curves is precisely at the  $g_k(t)$  mode.

The transmodal property of Theorem 1 is in the same spirit as Falmagne's (1968) fixed-point property of binary mixtures, which states that the density functions of all binary mixtures of the same processes must intersect at the same time point if they intersect at all. Just as the fixed-point property must hold for all binary mixtures, Theorem 1 must hold for all values of  $k$  for which pure insertion and the exponential assumption are appropriate, and thus a very ordered picture of the set of RT densities from the various experimental conditions emerges. This ordering is illustrated in Figure 2, again with gamma densities having the same rate, for the case in which  $k$  runs from two through five. The two-stage gamma is the leftmost curve, and the five-stage gamma is the rightmost. Between these, the ordering induced by the number of stages is preserved. In all cases, the Theorem 1 prediction that the  $g_k$  and  $g_{k-1}$  densities intersect at the mode of  $g_k(t)$  is upheld.

A question that naturally arises in connection with Theorem 1 is whether any models not making the two assumptions can also predict the Figure 2 result. Unfortunately, this possibility cannot be ruled out, and it is rather easy to see why not. Note that in the proof of Theorem 1, no use was made of the fact that  $V_k$  is a positive constant, but only that it is always positive. This admits the possibility that other representations could yield densities corresponding to this same pattern,

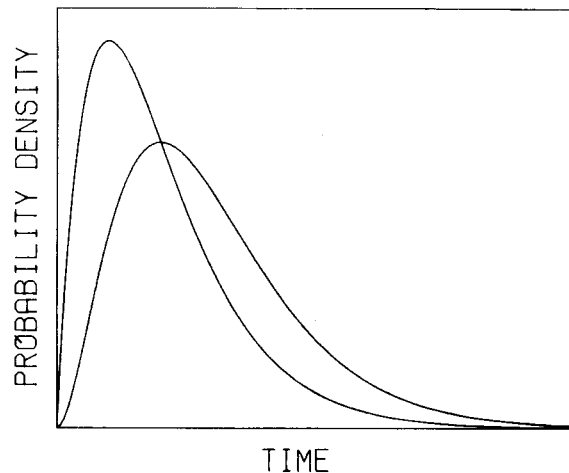


Figure 1. Two gamma densities. The leftmost curve is a two-stage gamma, and the rightmost curve is a three-stage gamma with the same rate. Theorem 1 predicts that they will intersect at the mode of the three-stage gamma.

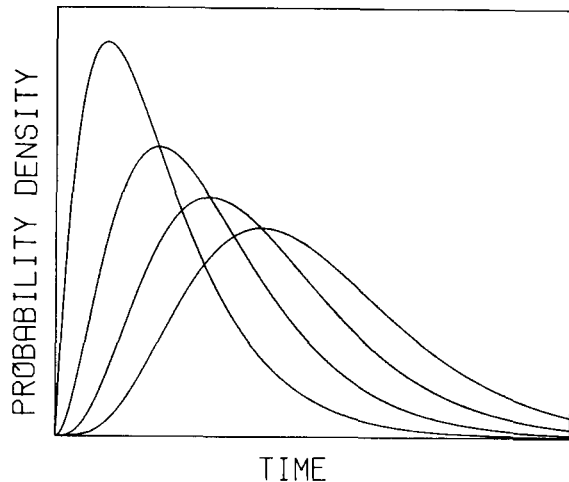


Figure 2. Four gamma densities, all with the same rate. The leftmost curve is a two-stage gamma, and the rightmost curve has five stages. Between these are three- and four-stage densities.

and so, unlike the Equation 4 test, Theorem 1 cannot be stated in an "if-and-only-if" fashion.

To review the development as it now stands, Theorem 1 and Equation 4 both provide tests of the double assumption of pure insertion at the distributional level and that the duration of the inserted stage is exponentially distributed (Theorem 1 through its falsification). The Equation 4 result is stronger, however, since it is both necessary and sufficient for the assumptions to hold, and since it also provides us with an estimate of the processing rate of the inserted stage. The primary advantage of Theorem 1 is the simplicity with which it can be invoked. It is a much more qualitative result. The empirical densities can be plotted and the condition can be checked. If it is obvious that the condition does not hold, then there is no need to invoke the more involved Equation 4 test, for the assumptions can be immediately rejected. On the other hand, if the conditions of Theorem 1 appear to be approximately satisfied, there is apparently no good reason to worry about developing a rigorous statistical test to decide whether they are. Instead, the more powerful Equation 4 test can be applied, for which some statistical decision rules are already available.

### An Empirical Application

For an empirical application of the test suggested by Theorem 1, I chose the same memory scanning data (i.e., from Townsend & Roos, 1973) employed by Ashby and Townsend (1980) in their application of the Equation 4 test. In a typical memory scanning experiment (see, e.g., Sternberg, 1966), a subject is first shown a short list of items (e.g., words or alphanumeric characters) called the memory set and then, a short time later, a single item called the probe. The subject's task is to respond "yes" or "no" as quickly as possible, depending on whether the probe was contained in the memory set. The idea is that adding an item to the

memory set might insert a stage into the RT process whose duration is the time it takes the subject to compare this new memory set item with the probe.<sup>2</sup>

The memory set in the Townsend and Roos (1973) experiment contained from one to five alphanumeric characters. Each of three subjects participated in 500 trials with each memory set size; 250 of these entailed "yes" responses (i.e., "probe is in the memory set") and 250 entailed "no" responses ("probe is not in the memory set"). Error rates averaged about 1.8%. The analyses reported below deal only with the target-absent, or "no," data. Here, the subscripted notation will come in handy;  $g_k(t)$  will refer to conditions in which there are  $k$  items in the memory set.

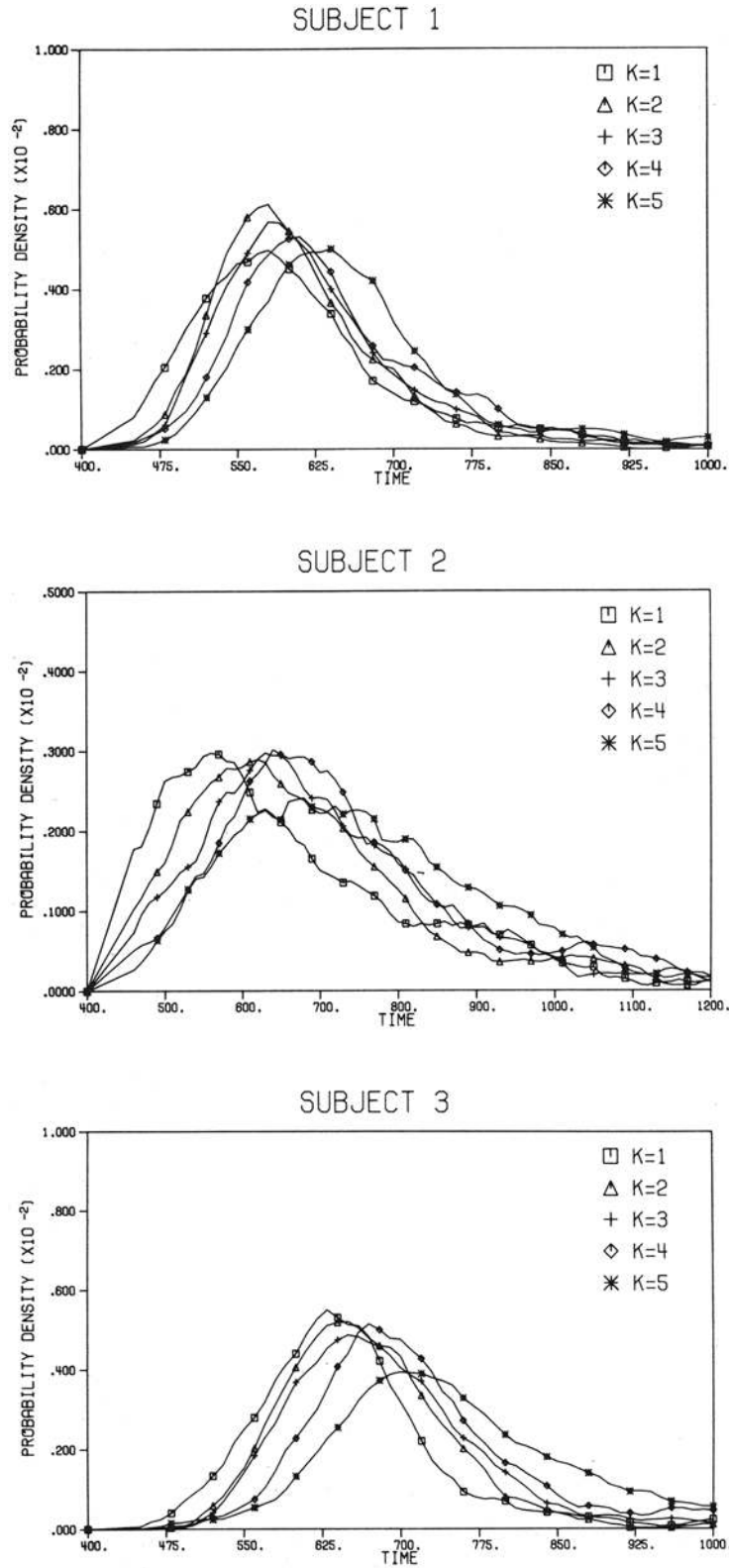
Since Theorem 1 is meant as a more qualitative result than Equation 4, it is recommended that the condition be checked by the eye rather than by some statistical procedure. This will be sufficient to detect gross violations. If there are none, the statistical decision rules developed for the Equation 4 test can be utilized. To aid the eye's task, some of the higher frequency components of the empirical densities were attenuated by passing the density estimates<sup>3</sup> through a Hamming window (see, e.g., Green & Luce, 1971) of width 100 msec for Subjects 1 and 3 and 120 msec for Subject 2, since the data of Subject 2 displayed more variability than did those of the others. This procedure will not change the mean of the density estimate, but it may change the mode slightly. On the other hand, the high-frequency components that the smoothing eliminates can make determination of the mode a difficult task.

Figure 3 shows the estimated RT densities for the three subjects for all five values of the memory set size. After some study, it can be seen that the condition of Theorem 1 is severely (i.e., obviously) violated for only one pair of densities, that is, for the pair  $k=1$  and  $k=2$  for Subject 1. For Subject 2, the  $k=5$  density has such a broad mode that it probably means little that  $g_4(t)$  passes through it. Overall, the assumptions of pure insertion at the distributional level and that the inserted stage duration is exponentially distributed can be summarily rejected in only one instance.<sup>4</sup> Even though this set of assumptions may eventually prove false, an appeal to the statistical test suggested by Equation 4 is clearly called for.

The results of this application agree nicely with Ashby and Townsend's (1980) application of the latter test to the same data. They tentatively accepted the null hypothesis of pure insertion at the distributional level and that the duration of the inserted stage is exponentially distributed in 9 of the 12 instances. Both analyses led to the rejection of these assumptions for the pair  $k=1$  and  $k=2$  for Subject 1.

### DROPPING THE EXPONENTIAL ASSUMPTION: TESTS OF PURE INSERTION

Clearly, the combination of exponentiality and pure insertion is strong enough to impose some fairly striking



**Figure 3. RT density estimates for the three subjects in the memory scanning experiment reported by Townsend and Roos (1973). The parameter  $k$  represents the size of the memory set on target-absent trials.**

constraints on the observable RT data, constraints that are easily tested. In this section, the consequences of dropping all distributional assumptions will be investigated by considering models that assume only pure insertion.

While this will tend to weaken possible tests, the assumption of pure insertion at the distributional level, by itself, still imparts enough regularity to data sets to generate testable predictions. For instance, it was noted earlier that the assumption implies an ordering on the means, so that a possible test is to check if the empirical means are ordered. That is, a possible decision rule is to accept pure insertion if  $\overline{RT}_k > \overline{RT}_{k-1}$  and if the ordering is reversed to reject it. Unfortunately, this test is very weak, in the sense that many other sets of assumptions predict a similar ordering on the means. In fact, the prediction of ordered means is weak even for pure insertion at the distributional level, since the assumption can be shown to predict much stronger orderings.

*Lemma 1:* If pure insertion holds at the level of the distribution, then  $G_{k-1}(t) \geq G_k(t)$  for all  $t > 0$ .

In addition to imparting an ordering on the mean RTs, pure insertion thus implies the stronger ordering on the distribution functions. The word "stronger" here refers to the fact that an ordering on distribution functions implies a mean ordering but that a mean ordering does not imply an ordering on the distribution functions (e.g., Townsend & Ashby, 1978).

Lemma 1 suggests that a test of pure insertion is to compare  $\hat{G}_k(t)$  and  $\hat{G}_{k-1}(t)$  and reject the assumption if  $\hat{G}_{k-1}(t)$  is not always greater than or equal to  $\hat{G}_k(t)$ . Such a comparison will be facilitated by a statistical test such as the Kolmogorov-Smirnov one-sided test. The Kolmogorov-Smirnov is a nonparametric test that uses the maximum deviation of the two empirical distribution functions as its test statistic, that is,

$$\max_t [\hat{G}_k(t) - \hat{G}_{k-1}(t)].$$

It was found to be a more powerful goodness-of-fit test than the more widely known chi-square test (e.g., Massey, 1951).

Even stronger than an ordering on the distribution functions is an ordering on the hazard functions. The hazard function,  $h_k(t)$ , gives the conditional probability density that processing will be completed in the next instant given that it is not yet completed, and the function is defined as  $h_k(t) = g_k(t) / [1 - G_k(t)]$ . An ordering on the hazard functions is very strong, since it implies a concomitant ordering on the distribution functions, but the reverse ordering need not hold (e.g., Townsend & Ashby, 1978). Theorem 2, therefore, provides a stronger test of pure insertion than Lemma 1.

*Theorem 2:* If pure insertion holds at the distributional level and if  $h_{k-1}(t)$  is nondecreasing in  $t$ , then  $h_{k-1}(t) \geq h_k(t)$  for all  $t > 0$ .

The requirement that  $h_{k-1}(t)$  be nondecreasing means that the conditional probability density that processing will be completed in the instant after time  $t$ , given that it has not been completed by that time, is a nondecreasing function of  $t$ . With the exponential distribution, this conditional probability is time invariant (and hence, nondecreasing), and it is just this property that was the target of Sternberg's (1975) criticisms. Many densities that would make good candidates for models of the time course of cognitive processes have increasing hazard functions (e.g., the convolution of a normal and an exponential, gamma, Rayleigh, etc.), and in addition, there is some empirical evidence in support of this assumption. Burbeck and Luce (Note 3) report that in a signal detection task the RT hazard function to a weak signal (thus ensuring a nonnegligible decision component) increases approximately monotonically to an asymptote. Thus, the requirement that  $h_{k-1}(t)$  is nondecreasing does not appear excessively severe. Of course, this is an empirical question rather than a theoretical one, and fortunately, tests for increasing hazard functions do exist (see, e.g., Barlow & Proschan, 1965).<sup>5</sup>

Hazard functions are starting to play a more important role in RT theorizing (e.g., Green & Luce, 1967; McGill & Gibbon, 1965; Thomas, 1971; Townsend & Ashby, 1978; Burbeck & Luce, Note 3), and in addition, the estimation problem is receiving attention from statisticians (e.g., Miller & Singpurwalla, 1977; Rice & Rosenblatt, 1976). Thus, application of a test such as the one suggested by Theorem 2 is becoming more and more feasible.

It should be emphasized that neither the Lemma 1 nor the Theorem 2 results are sufficient to conclude that pure insertion holds. There are other models that can predict these same orderings. Even so, as tests, both are much more powerful than a simple comparison of means.

A condition that is necessary and sufficient for pure insertion is that the function  $f_k(t)$  of Equation 2 be a true probability density function, that is, that it is always nonnegative, it encloses an area of one, and since the variable of interest is time, it equals zero for all negative  $t$ . As mentioned earlier, these facts have led to various attempts to develop a technique to test pure insertion and, if it is supported, to estimate the unobservable  $f_k(t)$ . Green (1971) and Green and Luce (1971) made use of the fact that Equation 2 implies that the Fourier transform of  $g_k(t)$  divided by the Fourier transform of  $g_{k-1}(t)$  equals the transform of  $f_k(t)$ . Thus, the unobservable density,  $f_k(t)$ , can be estimated by taking the inverse transform of this ratio. Applying this technique to data obtained in a standard signal detection task consistently led to  $f_k(t)$  estimates with regions of negative density, thus impugning the assumption of pure insertion in these experimental circumstances.

This procedure, however, is not without its difficulties. Classical density function estimates contain many high-frequency components, and the process of dividing Fourier transforms tends to magnify these to such an extent that the resulting  $f_k(t)$  estimates are swamped by noise. Thus, if classical density estimates are employed, considerable smoothing must be performed before the technique is feasible. The use of smoothing functions such as the Hamming window (used by Green, 1971, and Green & Luce, 1971) cause a loss of information and also artificially increase the variance of the density estimate and thus may introduce a subtle bias into the whole procedure.

An alternative approach, which might prove more successful, is to use the newer density estimates based on spline transformations (e.g., Bloxom, 1979; Wahba, 1976). Basically, these consist of several segments of truncated polynomials joined end to end and constrained to have continuous first- and second-order derivatives at all points at which different segments join. Since these very smooth estimates contain virtually no high-frequency components, they might reduce some of the statistical problems associated with this procedure.

Kohfeld et al. (Note 1) slightly modified Green's (1971) approach by using the digitalized z-transform rather than the Fourier transform. The z-transform is insensitive to high-frequency components, and thus this method requires no extra smoothing. Using data from a signal detection task, Kohfeld et al. estimated the probability density function of the inserted stage duration and convolved this estimate with  $\hat{g}_{k-1}(t)$ . As is predicted by pure insertion, they consistently found the result to correspond closely with the  $g_k(t)$  estimate.

A modified version of this general technique might also be used to simultaneously test the added assumption that the duration of the inserted stage is exponentially distributed. For instance, under the assumptions embodied in Equations 2 and 3, the exponential rate,  $V_k$ , can be estimated via a modified method of moments from the mean RTs as  $\hat{V}_k = 1/(\overline{RT}_k - \overline{RT}_{k-1})$ . An exponential probability density function with this rate can now be convolved with  $\hat{g}_{k-1}(t)$ , and the result can be compared with  $\hat{g}_k(t)$  to test if Equation 2 holds. This method works almost as well when other distributional assumptions are made. For instance, in the case in which the inserted stage duration has a gamma distribution, Ashby and Townsend (1980) provide estimates of its parameters in terms of the means and variances of  $RT_k$  and  $RT_{k-1}$ .

Of course, it is also possible to obtain nonparametric estimates of all moments of the unobservable  $T_k$  distribution [i.e., since  $E(T_k) = E(RT_k) - E(RT_{k-1})$ ,  $\text{var}(T_k) = \text{var}(RT_k) - \text{var}(RT_{k-1})$ , etc.]. Under certain general conditions, these can then be used to estimate the unobservable  $f_k(t)$  (e.g., Sternberg, Note 2). This procedure, however, requires accurate estimation of at

least some higher moments (e.g., the third and fourth in the Pearson system), and unfortunately, these are associated with very large standard errors (e.g., Ratcliff, 1979). Thus, this approach is not likely to play a central role in RT theory development until some of its rather severe statistical problems are solved.

With regard to empirical applications of the nonparametric tests suggested above, Townsend and Ashby (in press) report the results of applying the Lemma 1 test to the same Townsend and Roos (1973) data considered earlier. The test of  $G_{k-1}(t) \geq G_k(t)$  reported by Townsend and Ashby was motivated by a suggestion of Sternberg (Note 4) that the condition might be incorporated into a test of self-terminating vs. exhaustive search strategies. As a test of pure insertion, one would certainly expect the distribution functions to be ordered as predicted by Lemma 1 in light of the fairly successful application of Theorem 1 and the Equation 4 test to this same data set. This is precisely the result found by Townsend and Ashby. For each subject, 20 tests were performed, 10 for the target-absent data [i.e.,  $G_1(t) \geq G_2(t)$ ,  $G_1(t) \geq G_3(t)$ , ...,  $G_4(t) \geq G_5(t)$ ], and 10 on the target-present data. Of the total of 60 tests for the three subjects, 59 accepted the null hypothesis of  $G_i(t) \geq G_j(t)$  for  $j > i$ .<sup>6</sup> The single rejection of this hypothesis was in the target-present data for Subject 2 when  $k = 4$  and 5. As it turned out, there was a concomitant decrease in the mean RTs in this case.

Although, as mentioned earlier, the condition of ordered distribution functions is perhaps not the strongest possible test of pure insertion, the overwhelming support found for it by Townsend and Ashby (in press) assures that pure insertion at the distributional level is at least a good first approximation in the Townsend and Roos (1973) data.

### RELAXING PURE INSERTION: TESTS OF THE EXPONENTIAL ASSUMPTION

This last section briefly considers tests of the exponential assumption that do not require pure insertion to hold. Specifically, I will address the following question. Given an observable RT density function, how can one tell whether the duration of one or more of the component processes or stages comprising the cognitive process is exponentially distributed? This is a difficult problem, and the answer must be more tentative than those in the previous sections.

The following result generalizes the arguments of Luce and Green (1970) and McGill and Gibbon (1965).

*Theorem 3:* Suppose that the durations of all RT components have nondecreasing hazard functions and that the duration of one or more of these is exponentially distributed. Let  $V$  be the rate of the slowest of these exponential stages (i.e., all other exponential rates are greater than or equal to  $V$ ). Under these conditions, the observable RT density function,  $g(t)$ , will have an

exponential tail,  $W \exp(-Wt)$ , for some constant  $W \leq V$ , and thus the tail of  $-\ln[1 - G(t)]$  will be linear, with slope  $W$ .

Thus, if the duration of one or more RT components is exponentially distributed, then under fairly general conditions, the observable RT density function will have an exponential tail, the tail of the RT hazard function will be constant, and the tail of  $-\ln[1 - G(t)]$  will be linear [where, as before,  $G(t)$  is the RT cumulative distribution function]. One obvious problem associated with this test is that the result does not say how far out in the tail it is necessary to go before it becomes exponential. Also, while Theorem 3 generalizes the conditions under which to expect the RT density function to possess an exponential tail, unlike the results of Luce and Green (1970) and McGill and Gibbon (1965), it does not provide a method of estimating the exponential rate  $V$ .

### SUMMARY AND CONCLUSIONS

Several new tests of the assumptions of pure insertion at the distributional level and that the duration of the inserted stage is exponentially distributed were suggested. Two of these were applied to the data of a memory scanning experiment reported by Townsend and Roos (1973). In fact, altogether three separate tests of these assumptions were applied to this same data set. Ashby and Townsend (1980) applied the Equation 4 test, and applications of the Theorem 1 and the Lemma 1 tests were reported above. The results of applying these three different tests were highly consistent, and the assumptions being tested were supported in each case. While there is possibility of error in any statistical test, the converging evidence provided by the successful application of each of the three tests to the Townsend and Roos (1973) data suggests that pure insertion and the exponential assumption may provide at least a good first approximation in many standard memory scanning tasks.

Whether pure insertion is a valid assumption in other experimental paradigms is a question that must be answered case by case. It is hoped, however, that the results presented here will provide at least a first step toward making this a question that, perhaps for the first time, can be answered empirically.

### REFERENCE NOTES

1. Kohfeld, D. L., Santee, J. L., & Wallace, N. D. *Loudness and reaction time: II Identification of detection components at different intensities and frequencies*. Manuscript submitted for publication, 1981.
2. Sternberg, S. *Estimating the distribution of additive reaction time components*. Paper presented at the annual meeting of the Psychometric Society, 1964.
3. Burbeck, S. L., & Luce, R. D. *Evidence from auditory simple reaction times for both change and level detectors*. Unpublished manuscript, 1981.

4. Sternberg, S. *Evidence against self-terminating memory search from properties of RT distributions*. Paper presented at the annual meeting of the Psychonomic Society, St. Louis, November 1973.

### REFERENCES

- ASHBY, F. G., & TOWNSEND, J. T. Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, 1980, **21**, 93-123.
- ATKINSON, R. C., HOLMGREN, J. E., & JUOLA, J. F. Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, **6**, 321-326.
- BARLOW, R. E., & PROSCHAN, F. *Mathematical theory of reliability*. New York: Wiley, 1965.
- BLOXOM, B. Estimating an unobserved component of a serial response time model. *Psychometrika*, 1979, **44**, 473-484.
- BORING, E. G. *A history of experimental psychology*. New York: Appleton-Century-Crofts, 1950.
- BORING, E. G. *History, psychology, and science: Selected papers*. New York: Wiley, 1963.
- CHRISTIE, L. S., & LUCE, R. D. Decision structures and time relations in simple choice behavior. *Bulletin of Mathematical Biophysics*, 1956, **18**, 89-112.
- DONDERS, F. C. [On the speed of mental processes] In W. G. Koster (Ed. and trans.), *Attention and performance II*. Amsterdam: North-Holland, 1969.
- FALMAGNE, J. C. Note on a simple fixed point property of binary mixtures. *British Journal of Mathematical and Statistical Psychology*, 1968, **21**, 131-132.
- GREEN, D. M. Fourier analysis of reaction time data. *Behavior Research Methods & Instrumentation*, 1971, **3**, 121-125.
- GREEN, D. M., & LUCE, R. D. Detection of auditory signals presented at random times. *Perception & Psychophysics*, 1967, **2**, 441-449.
- GREEN, D. M., & LUCE, R. D. Detection of auditory signals presented at random times: III. *Perception & Psychophysics*, 1971, **9**, 257-268.
- HOHLE, R. Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, 1965, **69**, 382-386.
- HUNT, E. Mechanics of verbal ability. *Psychological Review*, 1978, **85**, 109-130.
- 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.
- MASSEY, F. J., JR. The Kolmogorov-Smirnov test for goodness of fit. *American Statistical Association Journal*, 1951, **46**, 68-78.
- 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., & 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*. NTIS No. AD-A040999/SST, 1977.
- PARZEN, E. *Modern probability theory and its applications*. New York: Wiley, 1960.
- PARZEN, E. On the estimation of a probability density function and the mode. *Annals of Mathematical Statistics*, 1962, **33**, 1065-1076.
- POSNER, M., & MITCHELL, R. Chronometric analysis of classification. *Psychological Review*, 1967, **74**, 392-409.



RATCLIFF, R. A theory of memory retrieval. *Psychological Review*, 1978, **85**, 59-108.

RATCLIFF, R. Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 1979, **86**, 446-461.

RATCLIFF, R., & MURDOCK, B. B., JR. Retrieval processes in recognition memory. *Psychological Review*, 1976, **86**, 190-214.

RESTLE, F. *Psychology of judgment and choice*. New York: Wiley, 1961.

RICE, J., & ROSENBLATT, M. Estimation of the log survivor function and hazard function. *Sankhyā: The Indian Journal of Statistics (Series A)*, 1976, **38**, 60-78.

SMITH, E. E. Choice reaction time: An analysis of the major theoretical position. *Psychological Bulletin*, 1968, **69**, 77-110.

SNODGRASS, J. G., LUCE, R. D., & GALANTER, E. Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, 1967, **75**, 1-17.

STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, **153**, 652-654.

STERNBERG, S. Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, 1975, **27**, 1-32.

SURWILLO, W. W. Developmental changes in the speed of information processing. *Journal of Psychology*, 1977, **96**, 97-102.

TAYLOR, D. A. Stage analysis of reaction time. *Psychological Bulletin*, 1976, **83**, 161-191.

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

TOWNSEND, J. T. Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 1972, **25**, 168-199.

TOWNSEND, J. T. Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, N.J: Erlbaum, 1974.

TOWNSEND, J. T. A stochastic theory of matching processes. *Journal of Mathematical Psychology*, 1976, **14**, 1-52.

TOWNSEND, J. T., & ASHBY, F. G. Methods of modeling capacity in simple processing systems. In N. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, N.J: Erlbaum, 1978.

TOWNSEND, J. T., & ASHBY, F. G. *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press, in press.

TOWNSEND, J. T., & ROOS, R. N. Search reaction time for single targets in multiletter stimuli with brief visual displays. *Memory & Cognition*, 1973, **1**, 319-332.

WAHBA, G. Histosplines with knots which are order statistics. *Journal of the Royal Statistical Society (Series B)*, 1976, **38**, 140-151.

NOTES

1. This assumption is not limited to serial systems. For example, many parallel models satisfy Equation 1 (see, e.g., Ashby & Townsend, 1980). In these cases,  $T_k$  does not represent the duration of the inserted stage but, instead, an inter-completion time (see, e.g., Townsend, 1974), that is, the random time between the successive completions of two stages. Alternatively,  $T_k$  can be thought of as the random time by which the total processing time is delayed in the more complex task as compared with the simpler task, thus allowing for overlap of processing stages. Thus the results given below are not restricted to serial systems but are valid for any of the large classes of models satisfying Equation 1.

2. It should be reemphasized here that the development does not require this comparison of the memory set items to the probe to be performed sequentially. As was explained in Foot-

note 1 (see also Ashby & Townsend, 1980), a simultaneous comparison might occur. All that is required is that the random time,  $T_k$ , by which  $RT_{k-1}$  is lengthened when a new item is added to the memory set, be independent of the random time  $RT_{k-1}$ .

3. As in Ashby and Townsend (1980), the RT density functions were estimated following Parzen (1962) with a gaussian kernel (see either of these papers for a more detailed discussion of the estimation procedure). Conventional histogram estimates, however, should be sufficiently sensitive to detect severe violations of the assumptions.

4. For those wishing a more rigorous criterion for rejecting the transmodal property, note that under the null hypothesis that the property holds,  $g_{k-1}(m_k) - g_k(m_k) = 0$ , where  $m_k$  is the mode of the  $g_k$  density. Following the development of Parzen (1962), it can be shown that under this null hypothesis and with the density estimates used in this application, the statistic  $\hat{g}_{k-1}(m_k) - \hat{g}_k(m_k)$  is asymptotically unbiased and normally distributed, and its variance is approximately

$$\text{var}[\hat{g}_{k-1}(m_k) - \hat{g}_k(m_k)] \doteq [g_k(m_k)/(10n\sqrt{\pi})],$$

where  $n$  is the size of the sample used to estimate  $g_k(t)$ . This last result can be used to obtain an estimate of the standard error of the vertical distance between the two density estimates at the  $g_k$  mode (i.e., at  $m_k$ ), under the null hypothesis that the transmodal property holds. The obtained vertical distances can then be compared with this standard error, and the transmodal property can be accepted or rejected depending on the outcome of this comparison. One problem with this strategy is that  $m_k$ , the true mode of  $g_k$ , is not known and must be estimated from  $\hat{g}_k(\hat{m}_k) = \max_t \hat{g}_k(t)$ .

In the present application, the sample size,  $n$ , was always 250, and  $g_k(m_k)$  can be roughly estimated from Figure 3 by  $\hat{g}_k(\hat{m}_k)$  (giving values of about .005 for Subjects 1 and 3 and .0025 for Subject 2). Using these values, the estimated standard error of  $\hat{g}_{k-1}(m_k) - \hat{g}_k(m_k)$  is about .00106 for Subjects 1 and 3 and about .00075 for Subject 2. In every case except one, the distance  $|\hat{g}_{k-1}(\hat{m}_k) - \hat{g}_k(\hat{m}_k)|$  is less than 1 standard error. The one pair of densities for which the vertical distance is greater than 1 standard error is the same Subject 1,  $k=1$  and  $k=2$  pair that our visual analysis decided is a violation of the transmodal property.

5. Besides this test, there are various restrictions that non-decreasing hazard functions impose on the associated cumulative distribution function,  $G_{k-1}(t)$ , which might be useful in deciding whether  $h_{k-1}(t)$  is nondecreasing, without ever having to estimate this function. The reader is referred to a standard text, such as Barlow and Proschan (1965), for an elaboration. One of the more useful is that if  $h_{k-1}(t)$  is nondecreasing, then  $\log[1 - G_{k-1}(t)]$  must be concave. If  $\log[1 - G_{k-1}(t)]$  is not concave, then  $h_{k-1}(t)$  cannot be nondecreasing. In fact, if  $h_{k-1}(t)$  is nonincreasing, then  $\log[1 - G_{k-1}(t)]$  must be convex.

6. Using a one-sided Kolmogorov-Smirnov test with  $\alpha = .05$ .

7. Of course,  $g_k(t)$  may also have inflection points at which the derivative equals zero, but which are neither maxima nor minima. Clearly, under the conditions of the theorem,  $g_k(t)$  and  $g_{k-1}(t)$  must intersect wherever  $(d/dt)g_k(t) = 0$ .

8. Note that the following equality is not necessarily exact, since  $h(t)$  may approach  $W$  without ever reaching it. Even so, for any  $\epsilon$  arbitrarily small, it will always be possible to find a time  $t_0 > 0$ , such that for all  $t > t_0$ ,  $W - h(t) < \epsilon$ . In other words, the fact that the equality may not be exact will make absolutely no difference, empirically.

APPENDIX

Proof of Theorem 1

Equation 4, which follows from Equations 2 and 3 (Ashby & Townsend, 1980), implies  $V_k[G_{k-1}(t) - G_k(t)] = g_k(t)$ , for

all  $t > 0$ . Differentiating both sides with respect to  $t$  yields  $V_k[g_{k-1}(t) - g_k(t)] = (d/dt)g_k(t)$ , for all  $t > 0$ . The right-hand side equals zero only at the maxima and minima<sup>7</sup> of  $g_k(t)$ , and the left-hand side equals zero only at the intersection of  $g_k(t)$  and  $g_{k-1}(t)$ . The equality of the two sides establishes the theorem, QED.

**Proof of Lemma 1**

Under the assumption of pure insertion (i.e., Equation 1, together with independence),  $G_k(t) = P(RT_k \leq t) = P(RT_{k-1} + T_k \leq t)$ . Thus,  $G_k(t) = P(RT_{k-1} + T_k \leq t) = P(RT_{k-1} \leq t - T_k) \leq P(RT_{k-1} \leq t) = G_{k-1}(t)$ . The inequality follows, since  $T_k \geq 0$ , QED.

**Proof of Theorem 2**

$$\begin{aligned}
 h_{k-1}(t) &= h_{k-1}(t) \frac{\int_0^t [1 - G_{k-1}(t-x)] f_k(x) dx}{\int_0^t [1 - G_{k-1}(t-x)] f_k(x) dx} \\
 &= \frac{\int_0^t h_{k-1}(t) [1 - G_{k-1}(t-x)] f_k(x) dx}{\int_0^t f_k(x) dx - \int_0^t G_{k-1}(t-x) f_k(x) dx} \\
 &= \frac{\int_0^t h_{k-1}(t) [1 - G_{k-1}(t-x)] f_k(x) dx}{F_k(t) - G_k(t)} \\
 &\geq \frac{\int_0^t h_{k-1}(t-x) [1 - G_{k-1}(t-x)] f_k(x) dx}{F_k(t) - G_k(t)}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\int_0^t \frac{g_{k-1}(t-x)}{1 - G_{k-1}(t-x)} [1 - G_{k-1}(t-x)] f_k(x) dx}{F_k(t) - G_k(t)} \\
 &= \frac{\int_0^t g_{k-1}(t-x) f_k(x) dx}{F_k(t) - G_k(t)} \\
 &= \frac{g_k(t)}{F_k(t) - G_k(t)} \\
 &\geq \frac{g_k(t)}{1 - G_k(t)} \\
 &= h_k(t).
 \end{aligned}$$

Therefore,  $h_{k-1}(t) \geq h_k(t)$  for all  $t > 0$ , QED.

**Proof of Theorem 3**

By a generalization of Theorem 2 above, if all component hazard functions are nondecreasing, then  $h(t)$  is dominated by each component hazard function. In particular,  $h(t) \leq V$  for all  $t > 0$ . Further, the hazard function of the convolution of densities with nondecreasing hazard functions is nondecreasing (Barlow & Proschan, 1965), and therefore, by hypothesis, the hazard function of  $g(t)$ ,  $h(t)$ , must have some asymptote  $W \leq V$ . Thus, there will exist some  $t_0 > 0$  such that<sup>8</sup>  $h(t) \doteq W$  for all  $t > t_0$ . Now  $h(t)$  determines the density function,  $g(t)$ , and therefore,  $g(t) \doteq W \exp(-Wt)$ , for all  $t > t_0$ . Similarly,  $G(t) = 1 - \exp(-Wt)$ , for all  $t > t_0$ , and thus,  $-\ln[1 - G(t)] = Wt$ , for all  $t > t_0$ , QED.

(Received for publication March 17, 1981;  
 revision accepted November 11, 1981.)