



## On measuring the minimum detection time: A simple reaction time study in the time estimation paradigm

Yung-Fong Hsu\*

National Taiwan University, Taiwan

Kornblum's time estimation paradigm, together with the so-called 'race model', provides an appealing alternative for measuring the 'cut-off' which separates 'true' reaction times from anticipatory reaction times. However, the model is not precise enough to reveal the relation between the signal intensity and the 'cut-off'. Accordingly, Kornblum's model is extended with an emphasis on the measure of the 'cut-off'. Another aspect of the extension is to use a parametric method to analyse the data. In particular, it is assumed that the time estimation-induced latency is gamma distributed and the signal-induced latency is Weibull distributed, with the latter shifted by the 'cut-off'. The rationale behind the parametric assumption is discussed. For illustrative purposes, two pieces of experimental work are presented. Since the core of the race model is the assumption of an independent race between the time estimation process and the detection process, the first experiment tests whether, for the same signal intensity, the signal-induced latency distribution is invariant across different time intervals; the second experiment tests whether, for the same time interval, the time estimation-induced latency distribution is invariant across different signal intensity conditions. The data from the second experiment are also used to test various parametric assumptions in the model, which include the signal effect on the 'cut-off'. The new model fits the data well.

### 1. Introduction

Cognitive processes are complex and normally unobservable. One of the commonly adopted measures for inferring the mental activity governing a cognitive process is the latency of the participant's response to the stimulus. That is, one can take the latency (or delay) between the signal presentation and the response initiation as a dependent variable. This latency is typically regarded as a random variable that reflects the dynamic activity of the mental organization. This holds even in cases where the stimulus is very simple in nature and the participant's task is limited to detecting the presence of the stimulus.

\* Correspondence should be addressed to Yung-Fong Hsu, Department of Psychology, National Taiwan University, Taiwan (e-mail: yfhsu@htu.edu.tw).

While the latency of the response is generally regarded as a good indication of the elementary mental operations, the underlying components of the latency are still not fully understood. Over the past years, psychologists have developed a more or less standard format for the experimental procedure, to which the name 'simple reaction time' is attached.

A standard simple reaction time trial starts with a warning signal (WS), followed by a foreperiod (FP), at the end of which the action signal (AS) is presented. The participant is required to react in a prescribed manner to the AS as quickly and as accurately as possible (Luce, 1986). The trial typically ends with the participant's response to the AS. The delay between the onset of the AS and the button press thus defines the *reaction time* (RT).

Most experimenters are interested in observing the latencies of *detection responses* triggered by the AS. However, the observed responses are practically always a mix of detection responses and *anticipatory responses* — responses triggered by the WS or some other signal-unrelated information. Evidence for this is the existence of *premature responses* — those responses which occur prior to the signal onset. In addition, the lack of premature responses does not necessarily eliminate the possibility that the observed RTs still include some proportion of anticipatory responses, and thus it is reasonable to assume that the observed RT distribution is mixed with responses not elicited by the stimulus. This intuition prompts many researchers to truncate their data sets by excluding very fast and very slow RTs (i.e. RTs falling outside a certain range) from the analysis. However, this truncation lacks a sound theoretical basis. As Ulrich and Miller (1994) pointed out, this procedure can introduce bias because some extreme but valid RTs might be excluded. Moreover, slow anticipatory responses may still remain in the truncated data.

There are two standard experimental methods that attempt to eliminate anticipatory responses in the simple RT situation (see Luce, 1986, for a review). The first is the *catch-trial procedure* where the experimenter uses a constant FP duration and introduces some proportion of *catch trials* on which no AS is presented. The participant is required to react to the AS as quickly as possible, but to avoid responding on catch trials. It is believed that the observed RT distribution still includes some anticipatory responses because the participant still makes some premature responses or presses the response button on catch trials.

The second method involves a variable FP. In this procedure, the experimenter uses random FP durations, which are typically uniformly distributed or exponentially distributed, with or without catch trials. The participant is required to react to the AS as quickly as possible, but to avoid making premature responses. When the distribution of the FP is uniform, the mean RT is a decreasing function of the FP duration (see Ollman & Billington, 1972), which has been interpreted as indicating that this procedure does not completely eliminate anticipatory responses. Using the exponentially distributed FP is a good alternative to reduce the proportion of anticipatory responses because the momentary tendency for the AS to appear is not affected by the time elapsed since the WS (Luce, 1986). However, a small number of premature responses are still observable.

In the following, I describe an alternative paradigm aimed at separating 'true' responses from anticipatory responses.

### **1.1. Time estimation: an alternative paradigm**

The structure of a trial in a *time estimation paradigm* is similar to that in a simple reaction time task. In both cases, a trial begins with a WS followed by an FP; at the end of the FP an AS is presented. On some trials the AS is omitted. The instructions to the

participant are very different, however. In a time estimation paradigm, the participant is required to press the response key as closely as possible to the length of a fixed time interval (TI), counted from the WS, if the AS is not presented. The participant is also instructed to react to the AS as quickly as possible, if it is presented. Since the length of the TI is typically set to be identical or very close to the length of the FP, the task can only be achieved by estimating the length of the TI on every trial, whether or not the AS is scheduled to be presented. Those trials in which the AS is scheduled to be presented will be called *signal trials*, the remaining trials being the *no-signal trials*. I describe herein two situations fitting this description.

In Ollman and Billington's (1972) experiment, there were two FPs (600 and 900 ms) and one TI (1,100 ms). One of the devices was a feedback lamp, operative only for the signal trials. The lamp would light when the participant failed to press the response button within 200 ms after the signal onset. The AS, an auditory tone burst, was set to be presented either at 600 ms or at 900 ms after the WS on the signal trials. For the no-signal trials, the instructions for the participant were to press the response key 1,100 ms after the WS. For the signal trials, the instructions were to press the same key at a time that would prevent the feedback lamp from lighting. No-signal and signal trials were intermixed. Thus, at the beginning of each trial, the participant had no idea whether to expect a signal or not. Accordingly, the participant was forced to establish a mental 'deadline', around 1,100 ms, for the keypress on each trial.

The latency of the participant's time estimation process on the TI (1,110 ms) had small variance. Moreover, the length of the TI was much longer than that of the short FP (600 ms) and was just a little longer than that of the long FP (900 ms). Accordingly, there were two cases to consider for the responses on the signal trials. The authors proposed that in the long FP case the participant's response was initiated by the outcome of a race between the *time estimation process*, triggered by the WS, and the *detection process*, triggered by the AS, whereas in the short FP case the participant's response was only initiated by the detection process.

The preceding paragraph suggests the basic idea of the so-called *deadline model*. The validity of the deadline model was confirmed indirectly by Ollman and Billington (1972) showing that the latency distribution induced by the detection process in the long FP condition is identical to that in the short FP condition.

Kornblum's (1973) time estimation paradigm was somewhat different from that of Ollman and Billington (1972). The AS was a neon light and the FP was 3 s. The instruction to the participant, applied to both the signal trial and the no-signal trials, was to press a key exactly at the end of the FP (i.e.  $TI = FP = 3$  s). On any signal trial, a response before 3 s suppressed the AS.

No-signal trials and signal trials were equiprobable and presented in random order. Thus, during the early stage of each trial where the AS had not been presented, the participant had no idea which trial he/she was in: it might have been a no-signal trial or it might have been a signal trial. However, whenever the AS was presented, the participant knew that his/her estimate of the time elapsed during the trial had exceeded 3 s and that he/she should press the key as quickly as possible. Therefore, for the purpose of data analysis, it makes sense to distinguish between three types of trial: (a) no-signal trial, (b) signal trial with the participant's response before the signal occurred, and (c) signal trial with the participant's response after the signal occurred. From the viewpoint of the participant, (a) and (b) were indiscriminable.

The AS-triggered RT is commonly decomposed additively into the *detection latency*, the latency for the detection process, and the *motor latency*, the latency for the motor

execution process (Luce, 1986). On the basis of this idea, Kornblum constructed a *race model* with four assumptions:

- K1. The latencies on the no-signal trials and the latencies on the signal trials pertaining to the time prior to the occurrence of the AS are both induced by the time estimation process.
- K2. On the signal trials with a response given after the signal occurs, the participant's motor execution process is triggered either by the detection process or by the time estimation process, whichever terminates first.
- K3. The motor latency is a constant and the same regardless of the source of the detection process and the time estimation process.
- K4. The time-estimation-process-induced latency and the detection-process-induced latency are independent random variables.

Notice that Kornblum did not specify whether the motor latency could vary with features of the stimulus (such as the intensity). Nevertheless, assumption K3 suggests that the motor latency does not depend on the intensity. The reason is the following. Consider running one participant in two separate conditions, one involving a strong AS and the other a weak AS, using the same FP duration. Then, according to assumption K3, the motor latency should be the same for the time-estimation-process-induced latency in both conditions. This indicates that it should be the same for the detection-process-induced latency in both conditions.

Let  $t_m$  represent the motor latency and, using boldface capital letters to represent random variables, let  $\mathbf{T}'_a$  denote the time course of the time estimation process,  $\mathbf{T}'_i$  the time course of the detection process,  $\mathbf{T}_a$  the observed RT on the no-signal trials, and  $\mathbf{T}_i$  the observed RT triggered by the AS. Based on assumptions K1, K2 and K3, Kornblum expressed the observed RT,  $\mathbf{T}$ , on the signal trials as

$$\begin{aligned} \mathbf{T} &= \min\{\mathbf{T}'_a, \text{FP} + \mathbf{T}'_i\} + t_m \\ &= \min\{\mathbf{T}'_a + t_m, \text{FP} + \mathbf{T}'_i + t_m\} \\ &= \min\{\mathbf{T}_a, \mathbf{T}_i\}, \end{aligned} \tag{1}$$

where all the random variables were non-negative and except for  $\mathbf{T}'_i$ , which was counted from the end of the FP, all were counted from the WS.<sup>1</sup>

The above expression is not yet a complete description for the random variable  $\mathbf{T}$ , since the race between  $\mathbf{T}_a$  and  $\mathbf{T}_i$  only occurs after the signal onset. That is, technically, equation (1) does not define  $\mathbf{T}$  on all points of the sample space. In fact, together with assumption K1, one can depict the behaviour of  $\mathbf{T}$  as follows:

$$\begin{aligned} \text{Prob}(\mathbf{T} < t) &= \text{Prob}[\mathbf{T} < t \ \& \ \mathbf{T} \leq (\text{FP} + t_m)] + \text{Prob}[\mathbf{T} < t \ \& \ \mathbf{T} > (\text{FP} + t_m)] \\ &= \begin{cases} \text{Prob}(\mathbf{T}_a < t), & \text{if } t \leq (\text{FP} + t_m), \\ \text{Prob}[\min\{\mathbf{T}_a, \mathbf{T}_i\} < t], & \text{if } t > (\text{FP} + t_m). \end{cases} \end{aligned} \tag{2}$$

<sup>1</sup> Both  $\mathbf{T}_a$  and  $\mathbf{T}_i$  are assumed to be continuous random variables.

The independence assumption in K4 suggests that some simple mathematical relations can be readily obtained from (2). Specifically in the case that  $t > (FP + t_m)$ , one can rewrite (2) as:

$$\begin{aligned} \text{Prob}(\mathbf{T} > t) &= \text{Prob}[\min\{\mathbf{T}_a, \mathbf{T}_i\} > t] \\ &= \text{Prob}(\mathbf{T}_a > t \& \mathbf{T}_i > t) \\ &= \text{Prob}(\mathbf{T}_a > t)\text{Prob}(\mathbf{T}_i > t), \quad \text{assuming independence,} \end{aligned} \quad (3)$$

which was also what Kornblum obtained. Let  $F$ ,  $F_a$  and  $F_i$  be the cumulative distribution functions of  $\mathbf{T}$ ,  $\mathbf{T}_a$ , and  $\mathbf{T}_i$ , respectively. Then Kornblum rewrote (3) as

$$1 - F(t) = (1 - F_a(t))(1 - F_i(t)),$$

and so

$$F_i(t) = \frac{F(t) - F_a(t)}{1 - F_a(t)}, \quad (4)$$

in which the two terms,  $F(t)$  and  $F_a(t)$  on the right, can be estimated from the RT data on the signal trials and the no-signal trials, respectively.

Kornblum's (1973) experimental results showed that the observed RTs on the no-signal trials had a mean of 3.042 s and a standard deviation of 0.23 s. Furthermore, by plotting the empirical cumulative RT curves on normal probability paper, Kornblum found that the shape of that RT curve fitted the normal distribution. On the basis of (4), he also plotted the induced  $F_i$  on normal probability paper. From that plot, it could be seen that the median of the RT triggered by the AS was about 180 ms after the signal onset, with a standard deviation of 27 ms. From the same plot, he also estimated the 'minimum detection time' (or the 'starting point') of the signal-initiated RTs to be about 110 ms. This minimum detection time was interpreted by him as the 'cut-off' to separate anticipatory RTs from 'true' RTs, under a given set of conditions.

## 1.2. Discussion

I have two comments on Kornblum's study. First, the status of the minimum detection time is not completely clear in his model. The interpretation of the minimum detection time is closely related to assumptions K2 and K3, which indicate that the motor latency is invariant with the signal intensity. This would suggest that the minimum detection time is also invariant with the signal intensity in Kornblum's model, unless a more explicit description of  $\mathbf{T}_i$  is provided. For instance, one can assume that the mass of  $\mathbf{T}_i$  is 0 in a positive interval between 0 and  $d_i$  and is positive above  $d_i$ , and the value of  $d_i$  varies with the signal intensity  $i$ . However, this point is not clear in Kornblum's paper.

Second, according to the context of Kornblum (1973), it appears that the estimated 110 ms minimum detection time was obtained by choosing the smallest value of  $t$  in his data so as to obtain the first non-negative  $F_i(t)$  based on (4). However, owing to sampling variability, fluctuation of estimates of  $F_a$  and  $F$  could make the estimate of  $F_i$  unrealistic or unstable,<sup>2</sup> for example being negative or an unreasonable estimate of the minimum

<sup>2</sup> It is also not clear how  $F_a$  was estimated in Kornblum's paper, because the RT histogram for the no-signal trials and that for the signal trials prior to the occurrence of the AS were not exactly identical in practice, although theoretically they both had the same shape (cf. Assumption K1).

detection time. Actually, the instability of the estimate of  $F_i$  was pointed out by Ollman and Billington (1972). To obtain the estimate of  $F_i$ , Ollman and Billington were forced to use the supremum of  $F_i(w)$  for any  $w$  not bigger than  $t$  to represent  $F_i(t)$ , in order to avoid non-monotonicity and negativity. However, this estimation method needs to be justified, since it may, for example, overestimate  $F_i(t)$  due to the possibility that one of the  $F_i(w)$  ( $w \leq t$ ) is overestimated.

The time estimation paradigm and the race model have been applied successfully to the studies of more complex cognitive processes. For instance, Meyer, Irwin, Osman, and Kounios (1988) extended the above ideas and developed a new speed-accuracy decomposition technique that was successful in distinguishing between discrete and continuous models of information processing (see, however, DeJong (1991) for championing 'intersensory facilitation' as a possible interpretation of the results). Moreover, the race model was employed with success in conjunction with the so-called 'countermanding procedure' in the study of simple reaction time (Ollman, 1973). Variants of the model and the procedure have also been used successfully to study other phenomena (see, for example, Logan & Cowan, 1984; Osman, Kornblum, & Meyer, 1986). Since none of those studies concerns the effect of the signal intensity, the role of the minimum detection time is yet to be clarified.

As an extension of Kornblum's (1973) work, the present study focuses on the development of a practical methodology for the analysis of the RT data in the time estimation paradigm, with a special emphasis on the measure of the minimum detection time. In particular, it is interesting to see whether there is a reliable effect of the signal intensity on the minimum detection time. With the same model-fitting approach, this paper also presents a theoretical study on the effect of the signal intensity on the signal-initiated RT.

The paper then reports some results from two experiments that were designed to test whether (i) for a given signal intensity, the signal-induced latency distribution is invariant across different time intervals (Experiment 1) or (ii) for a given time interval, the time-estimation-induced latency distribution is invariant across different signal intensity conditions (Experiment 2). The data from the two experiments are mainly used as a methodological demonstration for the use of the direct model-fitting approach to the analysis of the time estimation paradigm.

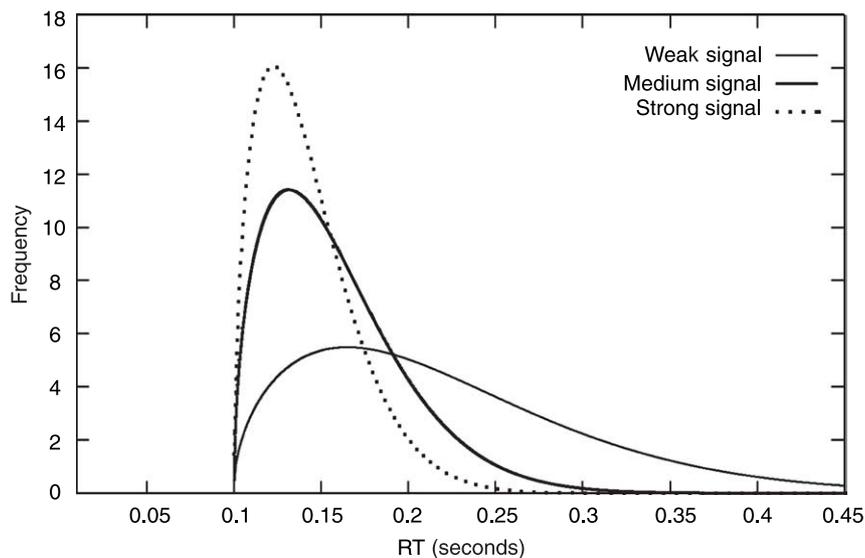
## 2. A modified model

Kornblum's time estimation paradigm, together with the race model, provide logical grounds for separating anticipatory RTs from 'true' RTs. In particular, Kornblum (1973) has shown the two (cumulative) empirical RT curves of the signal trials and the no-signal trials on normal probability paper. It is clear from this that there is a small region indicating the beginning of the deviation of the two curves. One point in this region was labelled by Kornblum as the minimum detection time. The role of the minimum detection time is not dealt with very effectively in the RT literature and, perhaps as a consequence, the relevant empirical estimation is not reliable. The following elaborates this issue.

It is generally agreed that as the signal intensity increases, the mean and variance of the RT random variable decrease. This can be seen from graphs that plot either the standard error of the mean of RTs (e.g. Mansfield, 1973) or the RT percentile (e.g. Dzhafarov, 1992; Smith, 1995) against signal intensity. It can be seen also from cumulative frequency polygons in which the RT curves from the weak signal conditions

are dominated by those from the strong signal conditions (e.g. Wandell, Ahumada, & Welsh, 1984).

However, the effect of the signal intensity on the lower tails of the RT histograms is not clear in those researchers' studies. Figures 1 and 2 illustrate this issue: they represent two almost identical sets of hypothetical signal-initiated RT distributions, except that in Fig. 2 those three RT distributions do not have the same minimum detection time as it is inversely related to the signal intensity. Obviously, in both cases the mean and variance of the RT random variable decrease as the signal intensity increases. Whether Fig. 1 or Fig. 2 is the more likely case is not clear from the above-cited researchers' RT data because their experimental paradigms, which are mostly of the standard format involving exponentially distributed FPs, cannot eliminate anticipatory RTs that partially contribute to the lower tails of the overall RT distributions. As a consequence, the minimum detection time of the signal-initiated RT has not been estimated reliably.

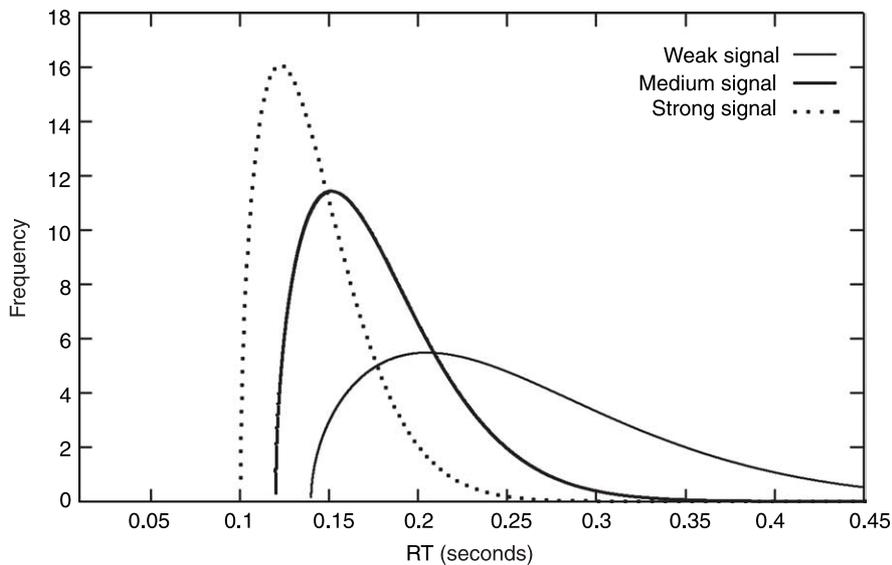


**Figure 1.** Three hypothetical RT distributions with the same minimum detection time (0.1 s after the signal onset).

The interpretation of the minimum detection time requires further investigation. In the following, the concept of the minimum detection time and its relation to the so-called 'irreducible minimum' are examined.

### 2.1. Three components of the RT process

The basic RT decomposition into decision latency and motor latency, briefly described in Section 1.1, is a simplification. The occurrence of the stimulus in an RT study is commonly regarded as initiating a stochastic process with three successive stages: the *predecision*, *decision* and *motor* processes (see Luce, 1986). The pre-decision latency is the time-required for the physical signal to be transduced from physical energy into neural spike trains plus the transit time for such pulses to pass from the output of the transducer to certain parts of the brain where the decision centre is located. This general



**Figure 2.** Three hypothetical RT distributions with three different minimum detection times (0.14, 0.12 and 0.1 s for the weak, medium and strong signals, respectively, after the signal onset).

scheme of RT decomposition is well supported. For example, a recent study by Zhang, Riehle, Requin, and Kornblum (1997) on monkeys has demonstrated that the time course of the neuronal activity in the primary motor cortex area is correlated with the time course of the psychological process responsible for the sensory, decision, and motor transformation in the RT task.

Physiological studies have shown that there are reliable effects of the signal intensity on neural processing at early levels of the visual system (see Nissen, 1977). In particular, it is known that the magnitude of the signal intensity can affect the amplitude of the neuron impulse, throughout the visual pathway, which in turn has an inverse effect on the response latency. For example, Vaughan, Costa, and Gilden (1966) studied the functional relation of the visual evoked response latency and the observed RT to different stimulus intensities. They found that the median of the visual evoked response latency, measured to the peak of the early positive deflection of the visual pathway, is between 76 and 100 ms for strong visual stimuli and between 190 and 220 ms for weak ones.

Physiological studies also have shown that when sensory neurons of the visual pathway are triggered by the signal intensity, the variance of the latencies of the initial impulse on those neurons is relatively small. For example, Levick (1973) used a microelectrode to measure the isolated impulses of cats' retinal ganglion cells. He showed that the magnitude of the latency variation in measures of the first  $n$  impulses ( $n$  is between 10 and 20) is fairly small, with standard deviation between 1 and 6 ms.

The decision process is assumed to be initiated by the arrival of the impulse in the decision centre and to terminate when the decision to respond is made. The latency of this decision process is commonly assumed to be a random variable, with both the mean and the variance decreasing as the signal intensity increases (see Luce, 1986).

After the decision stage, it takes some time for the motor neurons to complete the execution of the response. The latency of this motor process is usually assumed to be a constant or a random variable with very small variance (see Luce, 1986; Smith, 1995).

Recent studies by Ulrich and his colleagues (Ulrich, Leuthold, & Sommer, 1998; Ulrich, Rinkenauer, & Miller, 1998) suggest that there might be changes in the latency of the motor component as a function of the signal intensity.

### 2.2. Minimum detection time: Measure of constant delay of RT

It is conceivable that as the signal intensity increases, the sum of latencies of the three processes approaches a *limit random variable* which represents the case where the duration of nerve transmission time, transit time and motor time have been reduced to their physiological limits. If one assumes that the variance of the limit random variable is very small, then, without much loss of information, one can use the central tendency (i.e. mean or median) to represent it. This position is actually taken by most researchers. According to this view, there is a hypothetical asymptotic value of RT representing the limit random variable, called the *irreducible minimum* (see Luce, 1986; McGill, 1963; Piéron, 1952; Woodworth, 1938).

In practice, however, most observed RTs are contaminated by the latencies triggered by anticipatory responses that include some 'very fast' or even negative RTs. Therefore, the above intuition motivates some researchers to adopt the cut-off procedure in which a critical value (say, 100 ms) is chosen as a cut-off to exclude very fast RTs (Ulrich & Miller, 1994). However, this cut-off, which is constant across different experimental conditions, may not serve as a good criterion, because some RTs exceeding the cut-off may also be triggered by the anticipatory responses. In fact, the very existence of the effect of the signal intensity on the pre-decision latency also makes the choice of constant cut-off across different experimental conditions problematic.

On the basis of the above arguments, our understanding of the signal-initiated RT random variable is as follows. Since the main focus of this study is the effect of the signal intensity on the RT, it is assumed that other features of the stimulus are held constant. Let us denote by  $\mathbf{T}_i$  the random variable measuring the response latency to a stimulus of intensity  $i$ , where  $i$  is measured on a ratio scale and varies in a positive interval between, say,  $a$  and  $b$  ( $b > a$ , with  $b$  large). It can be assumed that the mass of  $\mathbf{T}_i$  is 0 for any values smaller than  $d_i$  ( $d_i > 0$ ). Thus,  $\mathbf{T}_i$  can be decomposed additively into a constant part  $d_i$  and a variable part  $\mathbf{X}_i$ :

$$\mathbf{T}_i = d_i + \mathbf{X}_i. \quad (5)$$

The function  $d : i \mapsto d_i$ , assumed to be strictly decreasing on the interval between  $a$  and  $b$ , corresponds to the minimum detection time for signal  $i$ . The variable part  $\mathbf{X}_i$  reflects mainly the decision process for the stimulus. Therefore, it also varies with the signal intensity.

Note that equation (5) is comparable to what was proposed in Kornblum's model, i.e.  $\mathbf{T}_i = \mathbf{T}'_i + t_m$  (see equation (1)). The subtle but important difference between the two equations is that in (5) the constant but signal-related component (one may call it  $c_i$ ) of  $\mathbf{T}'_i$  is spelled out and is added to the motor latency. That is,  $\mathbf{X}_i = \mathbf{T}'_i - c_i$  and  $d_i = t_m + c_i$ .

Equation (5) also suggests that the variable part measures the participant's ability to detect the AS, whereas the constant part measures the participant's physiological limit for the RT process. This postulate is consistent with the common observation that the mean and variance of observed RTs decrease as the signal intensity increases.

The concept of the irreducible minimum can thus be formalized as follows. Assuming that both the expectation and the variance of  $\mathbf{X}_i$  approach 0 as  $i \rightarrow b$ , then  $\text{Var}[\lim_{i \rightarrow b} \mathbf{T}_i] = 0$  and

$$E[\lim_{i \rightarrow b} \mathbf{T}_i] = E \lim_{i \rightarrow b} (d_i + \mathbf{X}_i) = d_b \approx \text{irreducible minimum.}$$

Very few studies have shed light on the behaviour of the minimum detection time. In the following, I give one example.

Ejima and Ohtani (1989) proposed a linear filter model for simple RT. In the model, there was a component equal to the sum of the irreducible and the absolute latency of the linear filter's response (see Ejima & Ohtani 1989, p. 123). This component is actually closely related to the proposed minimum detection time. Indeed, the estimated values of that component by their method showed reliable effects of the signal intensity. Unfortunately, they did not explicate this point and the status of the component was not clear.

### 2.3. Gamma-Weibull assumption

Under the time estimation paradigm, Kornblum's (1973) race model is a signal detection-type model for the reaction latencies, because 'true' RTs can be separated from 'mixed' RTs. While Kornblum intended his analysis to show how to estimate the minimum detection time under a given set of conditions, his race model was not precise enough to handle the effect of the signal intensity on the minimum detection time. Accordingly, in the following I extend Kornblum's work by proposing a modification of his race model.

Let  $d_i$ , indexed by the signal intensity  $i$ , be the constant delay (or the minimum detection time) of the RT process. Using the same notation as before, one adds  $d_i$  to the side conditions in equation (2):<sup>3</sup>

$$\text{Prob}(\mathbf{T} < t) = \begin{cases} \text{Prob}(\mathbf{T}_a < t), & \text{if } 0 < t \leq (\text{FP} + d_i), \\ \text{Prob}[\min\{\mathbf{T}_a, \mathbf{T}_i\} < t], & \text{if } t > (\text{FP} + d_i). \end{cases}$$

Furthermore, it is postulated that the time estimation RT is gamma distributed and the signal initiated RT is Weibull distributed but shifted by  $d_i$ , thus:

$$f_a(t) = \frac{\zeta^\eta t^{\eta-1} \exp(-\zeta t)}{\Gamma(\eta)}, \quad t > 0,$$

$$f_i(t) = m\lambda(t - d_i)^{m-1} \exp[-\lambda(t - d_i)^m], \quad t > d_i,$$

where 0 represents the onset of the WS and  $d_i$  is counted from the WS. Thus there are a total of five parameters in the model.

Both the gamma and Weibull distributions are appropriate for the study, because they have positive mass only for positive values and are very flexible. Nonetheless, the choice of the gamma distribution for the time estimation RT is not very critical. Earlier studies (e.g. Snodgrass, Luce, & Galanter, 1967) of the time estimation RT have shown that the distribution is roughly symmetrical, and so other distribution functions with the symmetry property, such as the normal which was implicitly assumed by Kornblum

<sup>3</sup> Note that the assertion of equation (1) (i.e.  $\mathbf{T} = \min\{\mathbf{T}_a, \mathbf{T}_i\}$ ) still holds without assuming that the motor latency is a constant. Also, despite the somewhat controversial defining boundaries of the pre-decision, decision and motor latencies, both the physiological and psychophysical studies seem to agree that the pre-decision latency and the motor latency are both random variables with negligible variances, whereas the decision process is the major source of the RT variability. Thus, it is very possible that  $d_i$  is mostly composed of the pre-decision latency and the motor latency.

(1973) in his analysis, are also good candidates for the fitting. It is possible to model the underlying time estimation process to obtain a suitable distributional assumption. However, since the distribution of the time estimation RT is not the focus of the current study, I do not attempt to model it.

Similarly, the Weibull distribution has suitable properties for the signal initiated RT in that it can have a long right tail, which is a distinct property of most RT distributions. In any case, a plausible motivation for the Weibull assumption is described in Section 3.2.

### 3. Power law of signal intensity on RT

The concept of the irreducible minimum has been discussed in Section 2.2. In fact, the study of the irreducible minimum is also closely related to the study of *Piéron's law*, which formulates the effect of the signal intensity  $i$  on the mean or the median RT as a power function (Luce, 1986; Piéron, 1952):

$$\text{mean or median RT}(i) = r + Ai^{-\alpha},$$

where  $r$ ,  $A$  and  $\alpha$  are positive parameters, with  $r$  the 'irreducible limit' (in Piéron's term), and  $A$  a dimensional constant. Hsu (in press) has studied this law extensively.<sup>4</sup>

Empirically, Piéron's law is known to hold for typical RT data obtained from the standard RT paradigm (Mansfield, 1973). It is of interest to investigate whether a similar form of Piéron's law holds for the signal-initiated RT derived from the time estimation paradigm. Specifically, according to equation (5) for signal intensity  $i$ , the RT random variable  $\mathbf{T}_i$  can be decomposed additively into a constant part  $d_i$  and a variable part  $\mathbf{X}_i$ :  $\mathbf{T}_i = d_i + \mathbf{X}_i$ . Taking, for example, the median (denoted by the symbol  $Me$ ) on both sides of the above equation, one gets  $Me(\mathbf{T}_i) = d_i + Me(\mathbf{X}_i)$ . The question of interest is whether  $Me(\mathbf{X}_i)$  can be expressed in a Piéron's law-type form:

$$Me(\mathbf{X}_i) = Bi^{-b}, \quad B > 0.$$

We shall call the above equation the *power law (for the signal-initiated RT)*.

In this section, it will be shown that assuming the signal-initiated RT to be Weibull distributed can lead to the above power law for the median RT, and, in fact, for all RT percentiles. Also, it will be shown that the shape parameter in the Weibull can partially determine the exponent parameter  $b$  in the power law.

#### 3.1. Weibull distribution assumption

The distribution of the shifted Weibull random variable  $\mathbf{W}$  has the form:

$$F(t) = \begin{cases} 0, & t \leq L, \\ 1 - \exp[-\lambda(t - L)^m], & t > L, \end{cases} \quad (6)$$

where  $\lambda$  is the scaling parameter,  $m$  the shape parameter, and  $L$  the location parameter. The mean and variance of the Weibull can be expressed as  $E(\mathbf{W}) = L + \lambda^{-1/m} \Gamma(1 + 1/m)$  and  $\text{Var}(\mathbf{W}) = \lambda^{-2/m} [\Gamma(1 + 2/m) - \Gamma^2(1 + 1/m)]$ , where  $\Gamma$  is the gamma function,  $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ .

<sup>4</sup> In particular assuming that both  $\alpha$  and  $A$  are functions of the background intensity and of the percentile rank, Hsu (in press) investigated from a theoretical viewpoint some possible functional forms of  $\alpha$  and  $A$ , with a special emphasis on the dependency of  $\alpha$  on the background intensity.

Proposing the Weibull (explicitly or implicitly) as one possible RT distribution is not new (e.g. Indow, 1993, 1995; Logan, 1992; Maloney & Wandell, 1984; Marley, 1989; McGill, 1963). In the following, I shall discuss some of these earlier applications.

Noting that for a steady light with a given intensity the number of active molecules in the photoreceptor increases linearly with time (so the hazard function is proportional to the time), McGill (1963, pp. 352–353) derived the density function of the latency to the first response of the photoreceptor when the stimulus is a steady light. The derived density function, which states that  $f(t) = nt \cdot \exp(-nt^2/2)$ , has a Weibull form (with  $L = 0$ ,  $m = 2$ ,  $\lambda = n/2$ ) although this point was not mentioned explicitly in his article.

Marley (1989) developed a large class of 'horse race' random utility models for choice probabilities and reaction times that can be represented in terms of multivariate (dependent) Weibull distributions. Later, Marley and Colonius (1992) discussed some subclasses of the above general class of models, such as those that satisfy the assumption of independence between the option chosen and the time of choice, and concluded that the above independence assumption is probably too strong for most data (see below).

As an exemplary case, Marley and Colonius (1992) (see also Marley, 1989) showed that the *separable* hazard function assumption, which also assumes independent functional forms of the choice probabilities and the reaction times, is consistent with that class of models. They noted that this separable hazard function assumption was proposed in Maloney and Wandell's (1984) simple RT study in the domain of visual detection. In that study, a special case of the separable hazard function assumption,  $b(ki, t) = \psi(k) \cdot b(i, t)$  (see also Luce, 1986, pp. 156–158), which implies that  $\psi(k) = k^\beta$ , with  $\beta > 0$ , i.e.  $b(i, t)$  is homogeneous of degree  $\beta$  (with respect to  $i$ ), was proposed. Marley and Colonius (1992) cited Luce's (1986, p. 156) study which showed that there is no multiplicative constant carrying one hazard function into another in Wandell *et al.*'s (1984) and other researchers' RT data, and concluded that the separability assumption is not 'entirely satisfactory'.

Let us have a close look at the form of the distribution function that Maloney and Wandell (1984) derived (see also Luce 1986, p. 157), namely

$$F(i, t) = 1 - \exp[-i^\beta \Lambda(t)], \quad (7)$$

where  $i$  is the signal intensity and  $\Lambda(t) = \int_0^t b(1, x) dx$ . Note that  $b(1, x)$ , the hazard function for unit intensity at time  $x$ , has an unknown functional form. Thus, it is not easy to verify the above equation empirically. Nevertheless, Wandell *et al.* (1984) noticed that for a given  $t$ , a straight line should be formed, with  $\beta$  the slope, by rearranging and taking logarithms twice on both sides of equation (7). Using this method, Wandell *et al.* (1984) validated the above equation and estimated the value of  $\beta$  at about 2.

Luce (1986, pp. 157–158) showed how the functional form of  $\Lambda(t)$  could have some impact on Piéron's law. Assuming that  $\Lambda(t) = ct$  (i.e. constant hazard, which means that the RT for unit intensity is exponentially distributed), he derived a power function relating the mean RT to the signal intensity, with the exponent parameter determined by negative  $\beta$  (i.e.  $-2$ ), a value much smaller than the generally accepted value (between  $-0.30$  and  $-0.35$ ).

Indow (1995), in his study of the asymptotic properties of the Weibull, used the Weibull to fit some existing data, including Wandell *et al.*'s (1984). Indow fitted those data with *composite* Weibulls, where the RT distribution is assumed to be composed of a few successive phases, with each phase Weibull distributed having different values of the parameters (see Indow, 1995, (II) of Fig. 12 and Table 1). Moreover, the location

parameter for the first phase of the composite Weibulls was questionably<sup>5</sup> set to 0 for the sake of simplifying the otherwise complicated estimation procedure. Nevertheless, using a simple Weibull to fit the data, one can still have a rough estimate about the shape parameter.

Note that the Weibull distribution (6) can be reformulated as a log-log equation by rearranging and taking logarithms twice on both sides; the shape parameter  $m$  will then be the slope in that equation. Fitting Wandell *et al.*'s (1984) data with the composite log-log plot of Weibull, Indow (1995) showed that the slope, which would be the estimate of the shape parameter had a simple Weibull been assumed, was much bigger than 1 (in fact, it was about 11; see Indow, 1995, (II) of Fig. 12 and Table 1). Taking this into account, one could generalize the functional form of  $\Lambda(t)$  proposed by Luce (see above) by assuming that the integral of the hazard function is a power function of the latency:  $\Lambda(t) = ct^q$ , with  $q > 1$ . As a result, the predicted reaction time distribution, based on (7), is a function of both  $i$  and  $t$  and has the Weibull form

$$F(i, t) = 1 - \exp(-i^\beta ct^q) = 1 - \exp[-c(i^{\beta/q}t)^q] = F(1, i^{\beta/q}t).$$

In other words, the scale of the distribution for intensity  $i$  is  $i^{-\beta/q}$  times the scale of the distribution for unit intensity. Consequently, the mean, standard deviation and RT percentiles should all decrease as a power function of the signal intensity,<sup>6</sup> with the shape parameter determined jointly by  $\beta$  and  $q$  (i.e.  $-\beta/q$ ). In the current example, it means the shape parameter would equal  $-2/11 = -0.18$ , a value much closer than  $-2$  (see above) to the generally accepted value (between  $-0.30$  and  $-0.35$ ).

### 3.2. Independent race of agents activated by intensity

Let us motivate a particular latency mechanism for the use of the Weibull as an RT distribution. Analogous to Logan's (1992) assumptions on the so-called 'instances of practice' in his study of automaticity, we make the following four assumptions:

- H1. Each signal intensity activates one or more (hypothetical) *agents*.
- H2. It takes each activated agent some time to deliver the information to the decision centre and to make a decision.
- H3. There is an independent race between the latencies of all activated agents. That is, as soon as one of the agent's activation levels reaches the decision criterion, the detection process is finished.
- H4. When the intensity increases, so does the number of active agents.

In the following, I specify the forms of some of the functions involved in these assumptions.

Assumption H4 is analogous to Logan's (1992) assumption that the number of 'instances' is proportional to the number of trials (practice), and here I assume a more general form where the number of (hypothetical) agents increases as a power function of the signal intensity. This generalization is appealing from the modelling viewpoint, since Hsu (2000) has shown that this power-function-induced Weibull form predicts the

<sup>5</sup> It is questionable because the cumulative RT curves in Wandell *et al.*'s (1984) study did not appear to start with the same time (see Wandell *et al.*, 1984, p. 650, Fig. 3)

<sup>6</sup> Note that the Weibull is not a necessary assumption to observe power function decreases in mean, standard deviation and RT percentiles (see Van Zandt & Ratcliff, 1995).

'power law' of RTS more accurately than does the linear-function-induced Weibull form.<sup>7</sup> Specifically, it is assumed that the number of agents  $g$  has the form

$$g(i) = \kappa i^\gamma, \quad (8)$$

where  $i$  is the signal intensity and  $\kappa, \gamma > 0$  are parameters. Regarding assumption H2, it is assumed that the latency for each activated agent is a random variable  $\mathbf{Y}$  and is Weibull distributed, with the location parameter equalling 0:  $F_{\mathbf{Y}}(t) = 1 - \exp(-\lambda t^m)$ .

Thus together with assumption H3, the latency of the detection process is the minimum of  $n$  independent and identically distributed (i.i.d.) Weibulls ( $n$ , depending on the signal intensity, is the number of the activated agents). Adding this assumption then forces the signal-initiated RT to be Weibull distributed<sup>8</sup> as well, with the same shape parameter (this is the so-called *min-stability* of the Weibull). Specifically, let the random variable  $\mathbf{W}_n$  (with the distribution function denoted by  $F_{\mathbf{W}_n}$ ) be the minimum of  $n$  i.i.d. Weibulls,  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ . One gets

$$\begin{aligned} 1 - F_{\mathbf{W}_n}(t) &= \text{Prob}(\mathbf{W}_n > t) = \text{Prob}(\min\{\mathbf{Y}_1, \dots, \mathbf{Y}_n\} > t) \\ &= \text{Prob}(\mathbf{Y}_1 > t \& \mathbf{Y}_2 > t \& \dots \& \mathbf{Y}_n > t) \\ &= \text{Prob}(\mathbf{Y}_1 > t) \text{Prob}(\mathbf{Y}_2 > t) \dots \text{Prob}(\mathbf{Y}_n > t) \quad (9) \\ &= [\exp(-\lambda t^m)]^n = \exp[-\lambda(n^{1/m}t)^m] \\ &= 1 - F_{\mathbf{Y}}(n^{1/m}t). \end{aligned}$$

From the above equation, one sees that the min-stability of the Weibull predicts a power function decrease (of the number of agents) on the random variable's scale:  $\mathbf{W}_n \stackrel{d}{\approx} n^{-1/m} \mathbf{Y}$ .

By simple algebra, it can be shown that the mean, standard deviation and all the RT percentiles follow the same power law of the signal intensity, with the exponent parameter in the power function (of the number of agents) the reciprocal of the shape parameter in the Weibull for a single activated agent. Based on equations (8) and (9), one sees that the corresponding exponent parameter in the power law for the signal initiated RT is  $\gamma/m$ :

$$F_{\mathbf{W}_n}(t) = F_{\mathbf{Y}}(g^{1/m}t) = F_{\mathbf{Y}}(\kappa i^\gamma)^{1/m}t) = F_{\mathbf{Y}}(i^{\gamma/m} \kappa^{1/m}t). \quad (10)$$

The validity of the model will be investigated analytically in Experiment 2 (see Section 5). Note, in passing, a consequence of equation (10): for intensities  $j$  and  $k$  ( $j \neq k$ ), with the estimated scaling parameters in the Weibulls being  $\lambda_j$  and  $\lambda_k$ , respectively, one gets

<sup>7</sup> In particular, by rearranging and taking logarithms twice on the Weibull form developed by Logan (1992), Hsu (2000) noticed that one could obtain a log-log function that is linear with respect to one of the (transformed) variables, with a slope of 1. However, Hsu's (2000) data analysis showed that most of Logan's (1992) data sets, when plotted in such way, showed a slope different from 1. On the other hand, the proposed more general form fitted Logan's (1992) data quite satisfactorily.

<sup>8</sup> Note that Logan (1992) obtained the Weibull form of the RT by an asymptotic argument (i.e. using the theory of extreme statistics) on the so-called 'instance'. However, this point has been shown to be not completely correct (see Colonius, 1995). Here I do not make such an asymptotic argument on the agent. On the other hand, as Cousineau, Goodman, and Shiffrin (2002) have shown, the i.i.d. assumption is not necessary to draw conclusions about the Weibull as the asymptotic distribution of minima.

$$\gamma = \frac{\ln(\lambda_j/\lambda_k)}{\ln(j/k)}, \tag{11}$$

which is empirically testable.

### 3.3. Fit of modified model to Kornblum’s data

The idea of using a parametric method to estimate the RT distribution in the time estimation paradigm is not new. Ollman and Billington (1972), when dealing with data obtained from a very similar time estimation paradigm (see Section 1.1), thought of using the method of maximum likelihood to estimate the underlying signal-initiated RT distribution and then carrying out a goodness-of-fit test by means of the chi-square (or  $G^2$ ) statistic. But they found the whole procedure ‘burdensome’ at that time (pp. 319–320).

Let us begin with the test of the gamma-Weibull model using Kornblum’s (1973) RT data.<sup>9</sup> Parameter estimates were sought so as to optimize the fit between the observed frequencies and the expected frequencies in the sense of maximum likelihood. To this end, the C version of the PRAXIS program is used<sup>10</sup> for the optimization.

The frequencies of the no-signal (signal) trials are divided into 53 (32) class intervals, with each bin 10 ms wide. One thus has 83 degrees of freedom in the data. Therefore, the degrees of freedom in the test amount to 78 (83 minus 5 estimated parameters). One obtains a  $G^2$  value of 55.9, which is considerably less than the 78 degrees of freedom ( $p = .97$ ), indicating an excellent fit.<sup>11</sup>

The results are summarized in Table 1, which also includes the results from Kornblum for comparison. It can be seen that the estimates of mean and standard deviation of the RTs on the no-signal trials are very close to what Kornblum estimated. Furthermore, one now obtains a more precise estimate of the minimum detection time  $d_i$ , which is about 136 ms after the onset of the signal. This value is somewhat larger than the 110 ms ‘cut-off’ that Kornblum estimated. It can also be concluded that the mean of the signal initiated RT is about 54 ms longer than the minimum detection time, with a standard deviation of about 29 ms.

**Table 1.** Comparison between estimated results from the modified model and those from Kornblum. Note that  $\mu_a$  is counted from the onset of the WS;  $d_i$  is counted from the onset of the AS;  $\mu_i$  is counted from the end of  $d_i$

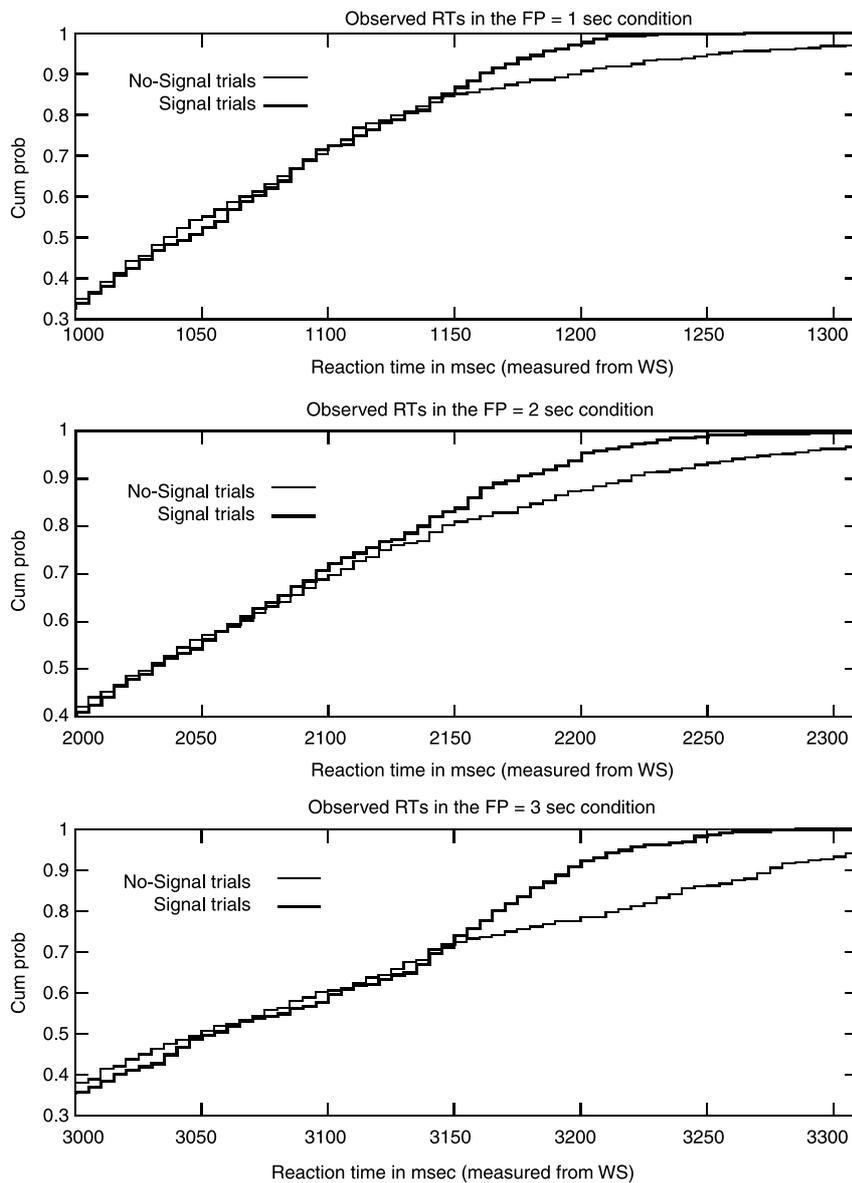
Models	$T_a$		$T_i$		$d_i$	$M$
	$\mu_a$	$\sigma_a$	$\mu_i$	$\sigma_i$		
Modified	3.044	0.23	0.054	0.029	0.136	1.89
Kornblum	3.042	0.23	N/A	0.027	0.110	

Note that the estimate of the shape parameter  $m$  in the Weibull is 1.89. However, since only one signal intensity was used in the experiment, at this stage it is difficult to

<sup>9</sup> The data were incomplete because only the RT histograms after the end of the FP are available (see Kornblum, 1973, p. 110).

<sup>10</sup> This program is an implementation of Brent’s (1973) algorithm for minimizing a multivariate function without using derivatives (Gegenfurtner, 1993).

<sup>11</sup> One might suspect that the fit is too good and might be an artefact. A close inspection of the analysis reveals that the overfit is mostly from the fit to the frequencies on the no-signal trials. Summing up the individual chi-square values, one obtains 34.6 for the no-signal trials (53 bins) and 20.8 for the signal trials (32 bins). This means the fit is actually not that dramatic when only focusing the analysis on the frequencies on the signal trials.



**Figure 3.** Observed cumulative RT curves on the no-signal trials and the signal trials in the FP = 1, 2, 3 s conditions of Experiment 1. Note that only the responses made after the end of FPs are shown.

pin down the relation between the shape parameter in the Weibull and the exponent parameter in the power law, as derived in the previous section.

The remainder of this paper describes two studies using the time estimation paradigm.

#### 4. Experiment 1: Three FPs with one signal intensity

The assumption of an independent race between the time estimation process and the detection process plays an important role in the race model, because it provides a simple

mechanism for separating signal-initiated RTs from 'mixed' RTs. The implication of the assumption suggests the following test: for the same signal, the signal-induced latency distribution should be invariant across different time intervals. If such an invariance is not found in the data, it would suggest that the independent race assumption needs to be modified. Experiment 1 is an illustration of such a design for the test.

#### 4.1. Method

##### 4.1.1. Apparatus

The apparatus included a 386 PC with a VGA card to generate the visual signals, a monitor to show the visual signals, and a buttonboard connected to the computer that the participants use to make their responses.

##### 4.1.2 Stimuli

The WS was a small 7 pixel by 7 pixel '+' sign in the centre of the screen that lasted 50 ms. It was followed by a prescribed FP, counted from the onset of the WS. At the end of the FP there was either a 10 pixel by 7 pixel solid rectangle shown in the centre of the screen for 50 ms or no AS was presented, depending on the type of trial and on when the participant pressed the key.

##### 4.1.3 Design and procedure

The design, similar to Kornblum's, consisted of one signal intensity and three different FPs. The participant was instructed to press a key at the estimated end of a certain FP (i.e.  $TI = FP$ ). There were two types of trials: (a) an AS was scheduled to be presented at the end of the FP and (b) no AS was scheduled to be presented at the end of the FP. These two types of trials occurred in random order, with equal probability. If the participant pressed the key before the end of the FP, the AS was not presented, even though it might have been scheduled to be presented.

The trial started at the onset of the WS and ended when the participant pressed the key and received feedback. The interval between the feedback and next onset of the WS was about 1.5 s. There were two kinds of auditory feedback, depending on when the participant made the response. If the participant responded before the end of the FP, then a 500 Hz tone was generated from the PC for 100 ms as feedback and the AS would not be presented. If there was no AS presented and the participant responded after the end of the FP, then a 1,000 Hz auditory tone was presented for 100 ms as feedback after the keypress. Note that in the case where the participant actually saw the AS (this would occur when the AS was scheduled to be presented and the participant did not press the key before the end of the FP), the AS itself served as (visual) feedback.

The procedure was applied in a block fashion to three different FP durations (1, 2 and 3 s) and one strong signal intensity (200 grey level), so there were a total of three conditions. For each condition, there were four sessions, which consisted of five blocks each, which in turn consisted of 48 trials.

#### 4.2. Results

One participant (the author) contributed a total of 2,880 observations, made up of 960 observations for each of the three FP conditions. The results are plotted in Fig. 3. Following the likelihood ratio procedure<sup>12</sup> and using all the data on the signal and the

---

<sup>12</sup> However, the (minus twice the log) likelihood function is not differentiable with respect to one of the parameters,  $d_r$ . Thus it is questionable whether the likelihood function is asymptotically distributed as a chi-square random variable. Simulation studies, not reported here, have shown that the chi-square test statistic is conservative.

no-signal trials, a test is performed on the modified model with a total of 15 parameters: two parameters for  $\mathbf{T}_a$  (a gamma) and three parameters for  $\mathbf{T}_i$  (a Weibull shifted by  $d_i$ ) for each FP condition. One obtains a  $G^2$  value of 263.2, with 259 degrees of freedom in the test ( $p = .42$ ), indicating that the model fits the data very well.

Since the signal intensity used in the experiment is the same for the three FP conditions, it is natural to test the submodel with a total of nine parameters: two parameters for  $\mathbf{T}_a$  (a gamma) in each of the three FP conditions, and three parameters for  $\mathbf{T}_i$  (a Weibull shifted by  $d_i$ ) which are hypothesized to be the same across the three FP conditions. The fit is very good based on the likelihood principle: the  $G^2$  value is about 4.2, with 6 degrees of freedom in the test ( $p = .64$ ). The results are summarized in Table 2, which also includes the observed means and standard deviations on the no-signal trials for comparison.

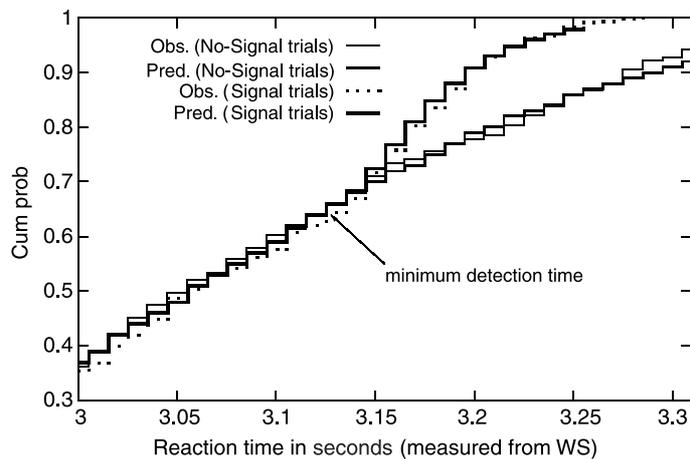
**Table 2.** Final estimates of the means and standard deviations of  $\mathbf{T}_a$  and  $\mathbf{T}_i$ , the minimum detection time ( $d_i$ ), and the shape parameter of the Weibull ( $m$ ) in each FP condition of Experiment I using a gamma-Weibull assumption. Note that  $\mu_a$  is counted from the onset of WS;  $d_i$  is counted from the onset of the AS;  $\mu_i$  is counted from the end of  $d_i$

Condition	$\mathbf{T}_a$		$\mathbf{T}_i$		$d_i$	$m$
	$\mu_a$	$\sigma_a$	$\mu_i$	$\sigma_i$		
FP = 1 s	1.044	0.108	0.067	0.044	0.129	1.54
Observed	1.047	0.116				
FP = 2 s	2.032	0.136	0.067	0.044	0.129	1.54
Observed	2.032	0.140				
FP = 3 s	3.056	0.177	0.067	0.044	0.129	1.54
Observed	3.048	0.177				

To give a clear comparison of the estimated RTs on the signal and the no-signal trials with the observed ones, Fig. 4 shows the predicted and empirical distributions for the 'FP = 3 s' condition. It can be seen that the four cumulative RT curves overlap considerably on the lower left-hand side of the figure, which is consistent with the interpretation that those RTs are all induced from the same time estimation process. After the minimum detection time, the curves for the signal trials start to deviate from those for the no-signal trials.

#### 4.3. Discussion

One might suggest that the assumption of independence of the time estimation process and the detection process may be tested by manipulating the proportion of the signal trials since, if the independence assumption is correct, then  $\mathbf{T}_i$  should remain the same in different signal presentation conditions. However, this argument is not necessarily correct, because manipulating the probability of the signal presentation can actually affect the participant's anticipation of the signal to be presented on the next trial. This mental anticipation can affect the participant's performance on the time estimate events. In other words, sequential effects occur. For example, the participant might have a strong bias towards waiting for the signal to be presented on each trial when the proportion of signal trials is high. After running some pilot studies in a block fashion, it was found that using the same number of scheduled signal trials and no-signal trials and



**Figure 4.** Comparison of the observed and predicted cumulative probabilities on the signal and no-signal trials for the FP = 3 s condition of Experiment 1.

mixing them at random is the best combination to reduce the possibility of the participant bias.

When the TI is short, the participant's time estimates are more accurate than when it is long. This can be seen from Table 2 where the standard deviation of the observed RTs on the no-signal trials increases with increasing FPs. This phenomenon is consistent with other investigators' findings (e.g. Snodgrass *et al.*, 1967). Considering the case where the TI is short and the FP is no earlier than the TI, one would see that most of the observed RTs are time estimation RTs. This indicates that one would not obtain much RT data resulting from the race between the time estimation process and the detection process. This in turn may affect the fit of the modified model (see also the remark by Kornblum, 1973). On the other hand, when the TI is long, the participant may have become impatient or unable to concentrate on the tasks. Consequently, the RT frequency histogram might be too flat and too variable to be analysed reliably using the likelihood ratio test. These two concerns are well taken into account by the design of the next experiment.

## 5. Experiment 2: Three signal intensities with one TI

Experiment 2 involves three ASs and one TI and is illustrated here as a possible design with three goals in mind: (i) testing whether the time-estimation-induced latency distribution is invariant across different signal intensity conditions; (ii) testing whether the minimum detection time is a decreasing function of the signal intensity; and (iii) testing whether the shape parameter of the Weibull distribution for the signal-initiated RT is the same for different signal intensity conditions.

The experimental procedure for Experiment 2 is a modification of the one used in Experiment 1.<sup>13</sup>

<sup>13</sup> Also, Experiment 2 was conducted using a different laboratory setting from that for Experiment 1.

## 5.1. Method

### 5.1.1. Apparatus

The participant was seated in front of a Leading Technologies monochromatic CRT display that was driven by an AT&T Truevision VISTA graphics card in a 486 PC. Only eight bits of the output of the graphics card were used for each pixel value. The 256 voltage levels associated with these eight bits were calibrated with a PR-650 Photo Research photometer to be nearly linear, with the lowest voltage level corresponding to 6 cd/m<sup>2</sup> and the highest one corresponding to 145 cd/m<sup>2</sup>. The participants responded by pressing a button on the buttonboard that was connected to the PC's keyboard Port.

### 5.1.2. Stimuli

The WS was a 7 pixel by 7 pixel ' + ' sign and the AS was a 10 pixel by 7 pixel solid rectangle shown in the centre of the monitor screen. It was a simple design, with only one TI (2 s) and three different ASs. The three ASs were generated by raising the intensities of the pixels in the target area to 10 grey level (about 11.5 cd/m<sup>2</sup>), 50 grey level (about 31.5 cd/m<sup>2</sup>) and 128 grey level (about 76.5 cd/m<sup>2</sup>), respectively, from the background surround, which was 0 grey level (about 6 cd/m<sup>2</sup>).

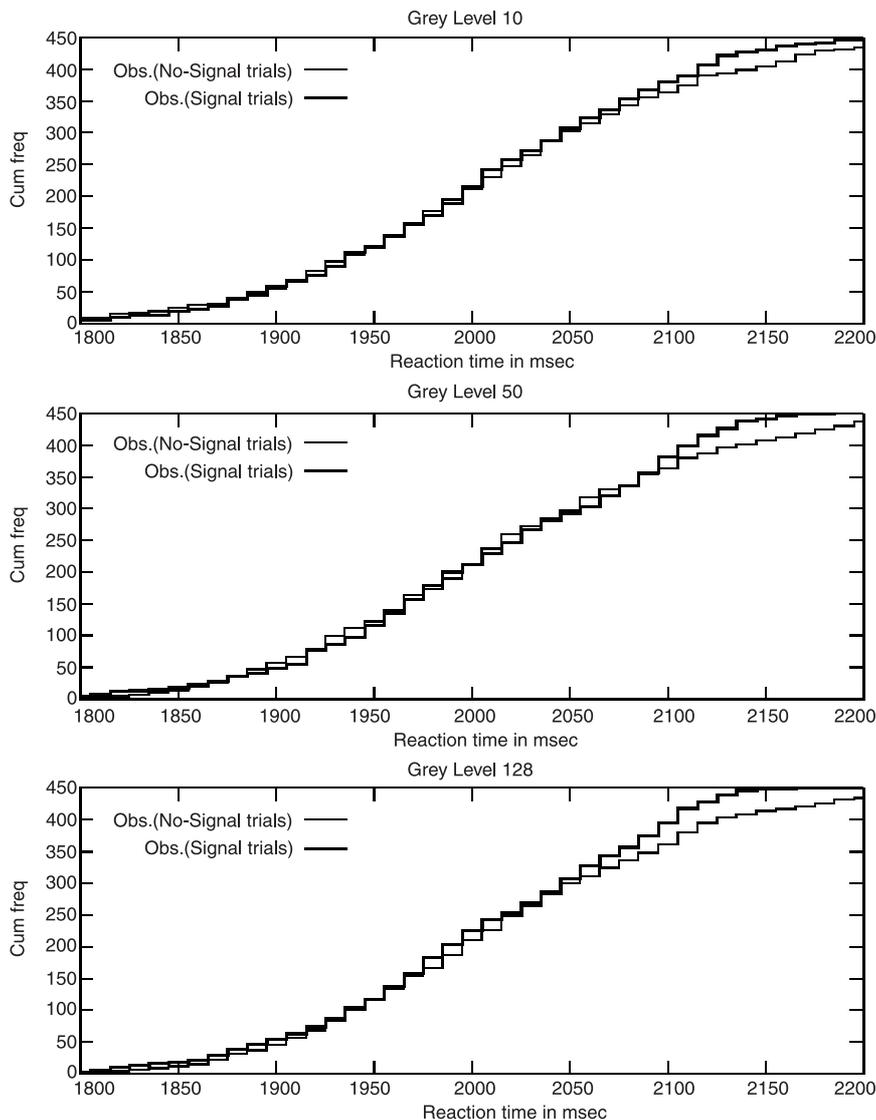
### 5.1.3. Design and procedure

The experimental procedure was the same as that Experiment 1, except for the following three minor changes: (1) the AS was response-terminated; (2) a 200 ms band surrounding the designated TI was introduced (if the participant pressed the key before 1,900 ms, then a 500 Hz tone was generated from the PC for 100 ms as feedback and the AS was not presented, even though it might have been scheduled to be presented; if there was no AS presented and the participant pressed the key after 2,100 ms, then a 2,000 Hz auditory tone was presented for 100 ms as feedback); (3) since the predecision process was mostly responsible for delaying the start of participant's detection process, the AS was scheduled to be presented a little earlier (1933.3 ms) than the assigned time for the time estimation task (2,000 ms) on the signal trials. Thus in the case where the participant actually saw the AS (this would occur when the AS was scheduled to be presented and the participant did not press the key before 1933.3 ms), the AS itself would more or less serve as unbiased feedback.

There were three sessions for each signal intensity condition. Each experimental session consisted of a warm-up block of 20 trials followed by five experimental blocks of 60 trials each, and so there were a total of 900 trials for each AS condition. Each block lasted about 5 minutes and there were short breaks between those blocks. The three AS conditions were run in a block-design fashion.

## 5.2. Results

One participant (not the author) was recruited for this experiment and the three different AS conditions were run at random. From the three graphs in Fig. 5, it can be seen that the data show a similar trend of departure of the RT curves on the signal trials from the RT curves on the no-signal trials after certain time intervals. Furthermore, the 'cut-off' indicating the start of the departure appears to occur earlier in the 128 grey-level condition than in other signal conditions, although the precise points cannot be determined by eye, due to sampling variability.



**Figure 5.** Observed cumulative RT curves on the no-signal trials and the signal trials in the 10 grey-level, 50 grey-level and 128 grey-level intensity conditions, respectively, of Experiment 2. Note that the participants' task is to estimate 2,000 ms but the signal onset is at 1933.3 ms.

Using the likelihood ratio procedure a test is performed on the modified model with two parameters for  $T_a$  (a gamma) and three parameters for  $T_i$  (a Weibull shifted by  $d_i$ ) for each AS condition. One obtains a  $G^2$  value of 194.2, with 207 degrees of freedom in the test ( $p = .73$ ). This indicates that the model with the gamma-Weibull assumption fits the data very well.

Note that the independent race assumption requires the time-estimation-induced latency distribution to be the same across the three AS conditions. Moreover, according to the theoretical analysis of the Weibull for the signal initiated RT (see Section 3), the shape parameter should be the same for different ASs. Thus, using the nested likelihood

principle, a test is also performed on the submodel with a total of nine parameters by further assuming that the shape parameter in  $\mathbf{T}_i$  and the two parameters in  $\mathbf{T}_a$  are the same across the three AS conditions. The  $G^2$  value is 4.4, with 6 degrees of freedom in the test ( $p = .62$ ), indicating a good fit. The results are summarized in Table 3.

**Table 3.** Final estimates of the means and standard deviations of  $\mathbf{T}_a$  and  $\mathbf{T}_i$ , the minimum detection time ( $d_i$ ), and the shape parameter of the Weibull ( $m$ ) in each signal intensity condition of Experiment 2 under the gamma-Weibull assumption. Note that  $\mu_a$  is counted from the onset of WVS;  $d_i$  is counted from the onset of the AS;  $\mu_i$  is counted from the end of  $d_i$

Condition	$\mathbf{T}_a$		$\mathbf{T}_i$		$d_i$	$m$
	$\mu_a$	$\sigma_a$	$\mu_i$	$\sigma_i$		
Weak AS	2.007	0.096	0.057	0.042	0.169	1.39
Observed	2.005	0.103				
Medium AS	2.007	0.096	0.047	0.035	0.143	1.39
Observed	2.006	0.098				
Strong AS	2.007	0.096	0.038	0.029	0.131	1.39
Observed	2.009	0.095				

From Table 3, one first notes that the estimates of the time estimation RTs are very close to the observed ones, which are almost identical in the three AS conditions. This result is consistent with the model.

The effect of the signal intensity can be seen from two sets of estimates. First, from Table 3, one sees that, with increasing signal intensities, the estimated values of mean and standard deviation of the signal-initiated RTs are decreasing. This phenomenon is consistent with the literature. Second, one also sees from Table 3 that the estimate of the minimum detection time is decreasing with increasing signal intensities. It strongly suggests that the 'cut-off' is signal-dependent. Notice that the estimated shape parameter is about 1.39. This value is not far from the obtained shape parameter in Experiment 1 (about 1.54).

With the shape parameter in the Weibull largely confirmed, let us study the relationship between the shape parameter in the Weibull and the exponent parameter in the power law. Since discriminability is basically determined by the signal contrast (see Smith, 1995), in the following, let us take the signal-to-background ratio as the index for the signal intensity. The weak, medium and strong signal-to-background ratios used for Experiment 2 were 1.9, 5.2 and 12.7, respectively. The estimated scaling parameters in those three signal intensity conditions are 44, 57 and 74, respectively. Taking any two intensity conditions and using equation (11), one obtains three estimates of the  $\gamma$  parameter, i.e. 0.25, 0.29 and 0.27, all fairly close estimations. The exponent parameter in the power function of the signal intensity can be easily obtained; it equals  $\gamma/m$ , which is between  $0.25/1.39 = 0.18$  and  $0.29/1.39 = 0.21$ .

In fact, the analysis based on the Weibull (see Section 3) suggests that both the mean and the standard deviation of the derived signal-initiated RT should follow a similar power law. Simple calculation on the estimated means of  $\mathbf{T}_i$  (see Table 3) shows that the estimated values of the exponent parameter are between 0.19 and 0.24. This result is consistent with the earlier estimation. A similar conclusion is reached when the

standard deviation is used. Thus, the assumption that a Piéron's law holds for the means, standard deviations and percentiles of the signal initiated RT is largely confirmed.

### 5.3. Discussion

The current (sub)model is consistent with Fig. 2. In fact, a competing model consistent with Fig. 1 can also be tested by constraining the minimum detection time to be the same over the three signal conditions, but allowing the two parameters of the Weibull to vary. A test was performed and the results showed that the estimated minimum detection time of the alternative model is about 139 ms, with the  $G^2$  value at about 8.2 ( $p = .22$ ). This value is somewhat higher than the  $G^2$  value of the current model (about 4.4,  $p = .62$ ). Based on the above results, it is concluded that the model of Fig. 2 is preferable to the model of Fig. 1.

The choice of the set of the parameters involved in the tasks is a challenge to the experimenter: both the time intervals used in the time estimation task and the ASs used in the detection task must be chosen appropriately for different participants. Failing to set the parameters properly would result in obtaining RT data that are not very useful in the analysis. In fact, according to Pins and Bonnet (1996), the photopic range of luminance is probably not the one corresponding to the largest changes of RTs (also see Mansfield, 1973). In Experiment 2, the intensity for the weak signal condition is about 11.5 cd/m<sup>2</sup>, still in the photopic range of luminance. However, the background intensity used in Experiment 2 (about 6 cd/m<sup>2</sup>) is also much brighter than those in other researchers' settings. A proper choice of setting for the background and the signal intensities is critical for future experiments.

## 6. Concluding remarks

In this paper, it has been shown that the time estimation paradigm can be used to measure the minimum detection time and the signal-initiated RT. A modified parametric race model was developed that assumes an independent race between the time estimation process (gamma distributed) and the detection process (Weibull distributed, but shifted by the minimum detection time). The validity of the model was largely confirmed using the nested likelihood ratio procedure. In particular, it was shown that the minimum detection time decreases as the signal intensity increases, and can be used therefore as a benchmark for the cut-offs used in the standard simple RT experiments in which signal intensity varies. Furthermore, if one obtains different estimates of minimum detection times for different participants, it might be the case that non-negligible individual differences occur at the stage of the sensory-motor process, which would invalidate the common exercise of averaging RT data from different participants.

The effect of the signal intensity on the signal-initiated RTs obtained from the time estimation paradigm has also been theoretically and empirically studied. A power law, which is a Piéron's-type law, was derived for the mean, standard deviation and all RT percentiles of the signal-initiated RT, with the exponent parameter in the power law determined by the shape parameter in the proposed Weibull for the (hypothetical) 'agent' activated by the signal intensity. Data analysis showed a great deal of agreement with these assumptions. Note that the current model should call for careful thought, in that other plausible alternatives might exist. Nonetheless, the derived signal-initiated RT distribution, based on the current model, could serve as a useful benchmark for the RTs obtained in the standard experiments. It might indicate the extent to which anticipatory

responses mix with detection responses, thus providing deeper insights into the strategy that participants might adopt in simple RT studies.

While it has been shown that the modified race model with a detection time component that exhibits Piéron's type law behaviour was not rejected, the work can only be summarized as a methodological demonstration. The time estimation paradigm needs to be further confirmed so as to make sure that it is appropriate for the measure of the minimum detection time and the signal-initiated RT. Thus, the findings need to be replicated with more participants. Moreover, signal intensities in the mesopic range of luminance should be studied in order to obtain more noticeable changes of RTs. The new data will contribute to clarifying (i) what the empirical laws of the minimum detection time and the signal-initiated RT are and (ii) what the estimated minimum detection time really measures.

There are also other ways to test indirectly what the minimum detection time measures. For instance, instead of using a finger press for the response, one could require the participant to use a foot press. Since the foot-pressing process is typically much slower than the finger-pressing process, the mean RT on the signal trials in the foot-pressing response mode will be larger than that in the finger-pressing response mode. Should the predecision latency solely dominate the minimum detection time, the estimated  $d_i$  in the two response modes will be very close to each other. Should the (constant) motor latency also contribute to the minimum detection time, the estimated  $d_i$  in the foot-pressing response mode will be considerably larger than that in the finger-pressing response mode. This topic is beyond the scope of the present study.

## References

- Brent, R. P. (1973). *Algorithms for minimization without derivatives*. Englewood Cliffs, NJ: Prentice Hall.
- Colonus, H. (1995). The instance theory of automaticity: Why the Weibull? *Psychological Review*, *102*, 744-750.
- Cousineau, D., Goodman, V., & Shiffrin, R. (2002). Extending statistics of extremes to distributions varying in position and scale and the implications for race models. *Journal of Mathematical Psychology*, *46*, 431-454.
- DeJong, R. (1991). Partial information or facilitation? Different interpretations of results from speed-accuracy decomposition. *Perception and Psychophysics*, *50*, 333-350.
- Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, *36*, 235-268.
- Ejima, Y., & Ohtani, Y. (1989). Analysis of simple reaction time to a sinusoidal grating by means of a linear filter model of the detection process. *Perception and Psychophysics*, *46*, 119-126.
- Gegenfurtner, K. (1993). Praxis: Brent's algorithm for function minimization. *Behavior Research Methods, Instruments and Computers*, *24*, 560-564.
- Hsu, Y.-F. (2000). A note on the power function decrease of the Weibull form in 'instance theory'. Manuscript.
- Hsu, Y.-F. (in press). A generalization of Piéron's law to include background intensity and latency distribution. *Journal of Mathematical Psychology*.
- Indow, T. (1993). Analyses of events counted on time-dimension: A soft model based on extreme statistics. *Behaviormetrika*, *20*, 109-124.
- Indow, T. (1995). *Weibull form in memory, reaction time, and social behavior: Asymptotic distribution of minima from heterogeneous population*. Technical Report Series, Institute for Mathematical Behavioral Sciences, University of California, Irvine, MBS 95-04.

- Kornblum, S. (1973). Simple reaction time as a race between signal detection and time estimation: A paradigm and model. *Perception and Psychophysics*, *13*, 108-112.
- Levick, W. R. (1973). Variation in the response latency of cat retinal ganglion cells. *Vision Research*, *13*, 837-853.
- Logan, G. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 883-914.
- Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*, 295-327.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Maloney, L. T., & Wandell, B. A. (1984). A model of a single visual channel's response to weak test lights. *Vision Research*, *24*, 633-640.
- Mansfield, R. J. W. (1973). Latency functions in human vision. *Vision Research*, *13*, 2219-2234.
- Marley, A. A. J. (1989). A random utility family that includes many of the 'classical' models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology*, *42*, 13-36.
- Marley, A. A. J., & Colonius, H. (1992). The 'horse race' random utility model for choice probabilities and reaction times, and its competing risks interpretation. *Journal of Mathematical Psychology*, *36*, 1-20.
- McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush & E. Galanter (Eds), *Handbook of mathematical psychology*, Vol. 1 (pp. 309-360). New York: Wiley.
- Meyer, D. E., Irwin, D., Osman, A., & Kounios, J. (1988). The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review*, *95*, 183-237.
- Nissen, M. (1977). Stimulus intensity and information processing. *Perception and Psychophysics*, *22*, 338-352.
- Ollman, R. T. (1973). Simple reactions with random countermanding of the go signal. In S. Kornblum (Ed.), *Attention and performance*, Vol. IV (pp. 571-581). New York: Academic Press.
- Ollman, R. T., & Billington, M. J. (1972). The deadline model for simple reaction times. *Cognitive Psychology*, *3*, 311-336.
- Osman, A., Kornblum, S., & Meyer, D. E. (1986). The point of no return in choice reaction time: Controlled and ballistic stages of response preparation. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 243-258.
- Piéron, H. (1952). *The sensations: Their functions, processes, and mechanisms*. New Haven, CT: Yale University Press.
- Pins, D., & Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Perception and Psychophysics*, *58*, 390-400.
- Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review*, *102*, 567-593.
- Snodgrass, J. G., Luce, R. D., & Galanter, E. (1967). Some experiments on simple and choice reaction time. *Journal of Experimental Psychology*, *75*, 1-17.
- Ulrich, R., Leuthold, H., & Sommer, W. (1998). Motor programming of response force and movement direction. *Psychophysiology*, *35*, 721-728.
- Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General*, *123*, 34-80.
- Ulrich, R., Rinkebauser, G., & Miller, J. (1998). Effects of stimulus duration and intensity on simple reaction time and response force. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 915-928.
- Van Zandt, T., & Ratcliff, R. (1995). Reaction time to incremental and decremental target luminance changes at various photopic background levels. *Psychonomic Bulletin and Review*, *2*, 20-54.

Copyright © The British Psychological Society

Reproduction in any form (including the internet) is prohibited without prior permission from the Society

284 Yung-Fong Hsu

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

Wandell, B. A., Ahumada, P., & Welsh, D. (1984). Reaction times to weak test lights. *Vision Research*, 24, 647-652.

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.

Zhang, J., Riehle, A., Requin, J., & Kornblum, S. (1997). Dynamics of single neuron activity in monkey primary motor cortex related to sensorimotor transformation. *Journal of Neuroscience*, 17, 2227-2246.

Received 16 July 2002; revised version received 20 September 2004