

# Time-dependent Poisson counter models of response latency in simple judgment

**Philip L. Smith\***

*University of Melbourne, Australia*

**Trisha Van Zandt**

*The Johns Hopkins University*

An important class of sequential-sampling models for response time (RT) assumes that evidence for competing response alternatives accrues in parallel and that a response is made when the evidence total for a particular response exceeds a criterion. One member of this class of models is the Poisson counter model, in which evidence accrues in unit increments and the waiting time between increments is exponentially distributed. This paper generalizes the counter model to allow the Poisson event rate to vary with time. General expressions are obtained for the RT distributions for the two- and the  $m$ -alternative cases. Closed-form expressions are obtained for response probabilities under a proportional-rates assumption and for mean RT under conditions in which the integrated event rate increases as an arbitrary power of time. An application in the area of early vision is described, in which the Poisson event rates are proportional to the outputs of sustained and transient channels.

## 1. Introduction

Response latency, or response time (RT), is a measure that occupies a position of central theoretical importance in many branches of sensory and cognitive psychology. In an effort to understand the processes underlying RT, mathematical psychologists have investigated models of performance in so-called 'simple judgment' tasks, that is, tasks that require a human observer to make a decision about some particular attribute of a stimulus display—such as magnitude, intensity, or numerosity—and to express this decision by making a speeded, typically binary, response. Of the various models that have been proposed for simple judgments, among the most influential have been the sequential-sampling models. These models represent the encoded stimulus attribute as a stochastic process and the decision mechanism as one in which successive values of the process are sampled and accrued until an evidence criterion is attained. From a theoretical perspective, the attraction of such models is that they provide a natural way to represent the processes that underlie RT and accuracy within a single, unified, theoretical framework. Detailed discussions of these models and their ramifications may be found in Vickers (1979), Townsend & Ashby (1983) and Luce (1986).

\* Requests for reprints should be addressed to Philip L. Smith, Department of Psychology, University of Melbourne, Parkville, Victoria, 3052, Australia (e-mail: philip@sherman.psych.unimelb.edu.au).

An important class of sequential-sampling models are those that have been variously termed *race* models, *parallel-counter* models, or *accumulator* models. These models assume that evidence for competing response alternatives accrues in parallel, as separate totals, with a response being emitted by the first to reach an evidence criterion. This contrasts with the other main class of sequential-sampling models, the random walks, and their continuous-time counterparts, the diffusion processes, in which evidence is accrued as a single, signed total, with a response being emitted when either an upper or lower criterion is reached (e.g., Link & Heath, 1975; Ratcliff, 1978). In models of either class, the time required for the accrual process or processes to first reach criterion is identified with the decision time component of RT. The proportion of correct responses is identified with the proportion of trials on which the criterion that is first attained is the one associated with the response corresponding to the stimulus presented.

The earliest parallel-counter model was proposed by LaBerge (1962), who, basing his development on the stimulus-sampling theory of Estes (1950), termed it a 'recruitment process'. A characterization of this model was given subsequently by Audley & Pike (1965) as one of a general class of Markov decision processes that included the random walk and the 'runs' model of Audley (1960). The recruitment model as originally proposed by LaBerge was formulated in discrete time and discrete state space. That is, at each of a set of discrete, equally spaced time steps, a unit increment was added to one of a pair of evidence counters. This had the undesirable consequence that, at low levels of stimulus discriminability, when the rates of accumulation in the two counters were similar, the model predicted RT distributions that were negatively skewed. These skewness properties arose as a consequence of the fact that, by virtue of the assumptions of discrete time and discrete state space, the sampling process was guaranteed to terminate within a fixed, maximum time.<sup>1</sup>

Subsequent authors attempted to preserve the essential structure of this model while modifying the assumptions in such a way as to obtain RT distributions that were better behaved. In practice, this involved dropping one of the assumptions of discrete time or discrete state space. For example, Vickers (1970, 1979) proposed an accumulator model in which evidence accrual occurred in discrete time but the increments to the evidence totals were continuously distributed. In Vickers's model, the increments to the evidence totals can be viewed as arising from the simultaneous action of a pair of half-wave rectifiers on a normal distribution of sensory effect. A mathematical characterization of this model was given by Smith & Vickers (1988); empirical tests and theoretical generalizations were described by Smith & Vickers (1989).

In an alternative approach, the unit-increment assumption is retained, but the interval between increments is assumed to be continuously distributed. Following Audley (1960), Pike (1966, 1973) suggested that the increment process could be represented as a Poisson process, resulting in exponentially distributed waiting times between increments. Pike (1973) also derived an expression for the relationship between the moments of the RT distributions

<sup>1</sup> The problem of negatively skewed RT distributions for the recruitment process only arises when one or other counter is incremented on every sample step. LaBerge (1962) also considered a version of the model with 'neutral elements', in which there is a fixed, non-zero probability that neither counter is incremented on any given step. Under these circumstances, the waiting time between increments is geometrically distributed, which approximates that of the Poisson model, in which the waiting time is exponential. The relationship between these versions of the model is discussed by LaBerge (1994).

for discrete-time and continuous-time models. The continuous-time version of the model was subsequently investigated in detail by Townsend & Ashby (1983), who represented the accrual process as a pair of independent Poisson processes, with unequal rates, one driving each of the evidence counters. Using this representation, they derived explicit expressions for the response-conditioned RT density functions, mean RTs, and choice probabilities, as a function of the Poisson rate parameters. This model has since been used to predict RT in models of spatial attention by LaBerge (1994) and Logan (1996). Subsequently, a detailed empirical evaluation of its performance was carried out by Van Zandt, Colonius & Proctor (2000) who showed that it successfully accounted for the RT distributions obtained in a same-different paradigm.

In this paper, we consider a further generalization of the parallel-counter model in which the Poisson rate parameters vary with time. The aim of this extension of the model is not to achieve the maximum generality possible (cf. Marley & Colonius, 1992), but rather to investigate its properties under a set of flexible but nevertheless constraining process assumptions that arise naturally in various applications. Stochastic accumulation mechanisms with time-varying accrual have been investigated in detail in a diffusion-process setting by Heath (1992) and Smith (1995, 1998, in press). In this paper, we investigate these properties in relation to the alternative class of parallel-counter models.

As discussed by Smith (1995, 1998, in press) there are a number of reasons why one may wish to consider models of information accrual with time-varying rates. Generally speaking, such models arise naturally when the assumptions associated with a particular domain lead to a model in which the outputs of stimulus encoding mechanisms are represented as non-stationary stochastic processes. For example, in the area of early vision, encoding mechanisms are often modelled as linear filters, whose impulse response functions depend on the spatial frequency properties of the stimulus (e.g., Watson, 1986). If the rate of information accrual is proportional to filter output, a model with time-varying rates naturally results.

Similar considerations apply when modelling performance on tasks that use brief stimulus exposures. If the observer's decision is assumed to be based on successive samples from a decaying perceptual trace, and the rate of decay during the sampling period is appreciable, then the stimulus encoding mechanisms are most naturally represented as non-stationary stochastic processes. Models of this kind also arise in connection with attentional paradigms, in which uncertainty about the spatiotemporal properties of the stimulus is manipulated. In this situation, a plausible model of performance is one in which the rate of information accrual is gated dynamically by selective attention. When combined with assumptions about the temporal properties of the gating function (e.g., Reeves & Sperling, 1986; Sperling & Weichelsgartner, 1995), a model with time-varying accrual rates again results. The mathematical machinery needed to represent such properties in a diffusion process setting was developed by Smith (1998, August; in press). The models described in this paper allow such properties to be expressed within an alternative, parallel-counter framework.

To this end, we consider a two-choice RT task in which one of two stimuli,  $S_A$  and  $S_B$ , is presented on any trial, and the observer responds either  $R_A$  or  $R_B$ . The model assumes that sensory encoding processes act on the stimulus to produce a pair of independent, parallel, point processes,  $X_A(t)$  and  $X_B(t)$ , whose event rates depend on time and on the stimulus presented. The 'events' in this setting may be thought of as the number of neural spikes occurring on a particular sensory channel, where events in  $X_A(t)$  represent evidence for response  $R_A$  and events in  $X_B(t)$  represent evidence for response  $R_B$ . We denote by  $a(t)$  and

$b(t)$ , respectively, the event rates on the  $R_A$  and  $R_B$  channels. For subsequent convenience, we also define the *integrated event rates*  $A(t)$  and  $B(t)$  as

$$A(t) = \int_0^t a(s) ds$$

and

$$B(t) = \int_0^t b(s) ds,$$

with the understanding that, in general,  $A(t) \geq B(t)$  when  $S_A$  is presented and  $B(t) \geq A(t)$  when  $S_B$  is presented. However, for simplicity, the dependence of event rates on the identity of the stimulus will be omitted from the notation.

Associated with the point processes  $X_A(t)$  and  $X_B(t)$  are a pair of counting processes,  $N_A(t)$  and  $N_B(t)$ . These are such that  $N_A(0) = 0$ ,  $N_B(0) = 0$ , and  $N_A(t)$  and  $N_B(t)$  equal, respectively, the number of point events that occur in the processes  $X_A(t)$  and  $X_B(t)$  during the interval  $[0, t]$ . With each of the counting processes we associate a *decision criterion*, denoted  $K_A$  and  $K_B$ , respectively. We then define the random variables  $T$ ,  $T_A$ , and  $T_B$  as follows:

$$\begin{aligned} T_A &= \inf\{t : N_A(t) \geq K_A\}, \\ T_B &= \inf\{t : N_B(t) \geq K_B\}, \\ T &= \min(T_A, T_B). \end{aligned}$$

That is,  $T_A$  is the time at which the evidence total  $N_A(t)$  first exceeds the criterion  $K_A$ ;  $T_B$  is the time at which the total  $N_B(t)$  first exceeds  $K_B$ , and  $T$  is the minimum of the two times. In the psychological model, the value of  $T$  is identified with the decision time component of RT on a given trial. The probability of the response  $R_A$ , denoted  $P(A)$ , is defined as

$$P(A) = P[T_A < T_B],$$

where implicit in this definition is the assumption that at least one of  $T_A$  and  $T_B$  is finite. As shown subsequently, a sufficient condition for the finiteness of  $T_A$  and  $T_B$  is that the associated integrated event functions  $A(t)$  and  $B(t)$  grow unboundedly in  $t$ . The extension of the model to the general  $m$ -alternative case is straightforward and is considered later in this paper.

We wish to obtain the joint RT density functions  $g_A(t)$  and  $g_B(t)$ , where these are defined as

$$\begin{aligned} g_A(t) &= \frac{d}{dt} P[T \leq t, T_A < T_B], \\ g_B(t) &= \frac{d}{dt} P[T \leq t, T_B < T_A]. \end{aligned}$$

The derivation proceeds in three stages. First, we obtain the properties of the time-dependent Poisson distribution. Second, we exploit the well-known duality between the Poisson distribution and the gamma waiting time distribution to obtain the probability density function (pdf) of the time-dependent gamma distribution. Finally, we use the standard properties of independent random variables to obtain the joint densities  $g_A(t)$  and  $g_B(t)$ . Although closed-form expressions for choice probabilities and mean RTs cannot be obtained in the general case, values of these statistics may be calculated quite straightforwardly from a knowledge of the joint density functions by numerical integration. Later in the paper we consider some special cases in which closed-form expressions for choice probabilities and/or mean RTs can be obtained. Finally, we consider an application in the area of early vision, in

which the Poisson event rates are proportional to the outputs of sustained and transient channels.

### 2. The time-varying Poisson process

As noted by Cox & Miller (1965) and Ross (1983), the derivation of the standard (time-homogeneous) Poisson process can readily be generalized to allow the event rate to depend on time.<sup>2</sup> To do so, we consider a point process,  $X(t)$ , and the associated counting process,  $N(t)$ , both for the moment unsubscripted. We stipulate that  $X(t)$  is such that events arrive with rate  $a(t)$  and that the numbers of events in non-overlapping intervals,  $(t_i, t_j]$ ,  $(t_j, t_k]$ , are independent. We require of  $a(t)$  only that it be positive and integrable.

To make the concept of ‘event rate’ precise, the following properties are ascribed to  $N(t)$ :

$$\begin{aligned} P[N(0) = 0] &= 1, \\ P[N(t + h) < n | N(t) = n] &= 0, \\ P[N(t + h) = n | N(t) = n] &= 1 - a(t)h + o(h), \\ P[N(t + h) = n + 1 | N(t) = n] &= a(t)h + o(h), \\ P[N(t + h) > n + 1 | N(t) = n] &= o(h), \end{aligned}$$

where, as usual,  $o(h)$  denotes a quantity of the order  $\lim_{h \rightarrow 0} o(h)/h = 0$ . In these equations, the independent-increments property is expressed by the fact that the right-hand sides depend on the interval  $h$  but not on the value of  $N(t)$ . Together with the law of total probability they imply that

$$\begin{aligned} P[N(t + h) = n] &= P[N(t + h) = n | N(t) = n]P[N(t) = n] \\ &\quad + P[N(t + h) = n | N(t) = n - 1]P[N(t) = n - 1] + o(h) \\ &= P[N(t) = n]\{1 - a(t)h\} + P[N(t) = n - 1]a(t)h + o(h). \end{aligned}$$

Letting  $P_n(t)$  denote the probability  $P[N(t) = n]$ , the preceding equation may be rearranged to give

$$\frac{P_n(t + h) - P_n(t)}{h} = -a(t)P_n(t) + a(t)P_{n-1}(t) + \frac{o(h)}{h},$$

which in the limit,  $h \rightarrow 0$ , yields the set of ordinary differential equations

$$P'_n(t) = -a(t)P_n(t) + a(t)P_{n-1}(t), \quad n = 0, 1, 2, \dots \tag{1}$$

Because  $P_{-1}(t) = 0$ , the equation for  $P_0(t)$  is homogeneous and so may be solved by simple separation of variables for the initial condition  $P_0(0) = 1$ . The remaining inhomogeneous equations for  $n = 1, 2, \dots$  may then be solved in succession by introducing the integrating factor  $\exp[-\int_0^t a(s) ds]$ , which renders the differentials exact. The relevant initial condition in each instance is  $P_n(0) = 0$ .

<sup>2</sup> Although the time-inhomogeneous generalization of the Poisson process is straightforward, it does not appear in most standard texts on stochastic processes. For this reason and for completeness, an outline of the derivation is provided here.

The general solution of the set of equations in (1) may then be expressed in terms of the integrated event rate function  $A(t)$  as

$$P_n(t) = \frac{[A(t)]^n e^{-A(t)}}{n!}. \quad (2)$$

In counter models of RT, the quantity  $P_n(t)$  in (2) represents the probability that the number of neural spikes arriving on a given sensory pathway during the sampling interval  $[0, t]$  is exactly equal to  $n$ .

To obtain the associated waiting-time distribution, we denote by  $T_n$  the quantity

$$T_n = \inf\{t : N(t) = n\}.$$

Our interest is in  $F(t)$ , the distribution function of  $T_n$  under the assumptions about the properties of the counting process  $N(t)$  stated previously. We denote the associated pdf by  $f(t)$  and the survivor function by  $\bar{F}(t)$ , where  $\bar{F}(t) = 1 - F(t)$ . From the duality of counting processes and waiting times we have  $\{N(t) < n\} = \{T_n > t\}$ , or, in a more explicit notation,  $\{\omega : N(t, \omega) < n\} = \{\omega : T_n(\omega) > t\}$  for all  $\omega \in \Omega$ , where  $\Omega$  is the underlying probability space, that is, the space of random sequences of point events. The equality signs in these expressions are to be understood in the sense of set equality and thus imply equality of measure. From this equality we obtain

$$\begin{aligned} P[N(t) < n] &= \sum_{j=0}^{n-1} P_j(t) \\ &= P[T_n > t] = \bar{F}(t). \end{aligned}$$

Summing terms of the form of equation (2) therefore yields the survivor function

$$\bar{F}(t) = \sum_{j=0}^{n-1} \frac{[A(t)]^j}{j!} e^{-A(t)}. \quad (3)$$

From this result we have the following lemma.

*Lemma 2.1.* Let  $N(t)$  be a time-varying Poisson counting process with integrated event rate  $A(t)$ . Then  $f(t)$ , the pdf of the waiting time required to obtain a total of  $n$  events is

$$f(t) = \frac{[A(t)]^{(n-1)} A'(t) e^{-A(t)}}{(n-1)!}. \quad (4)$$

We refer to this pdf as the *time-varying* or *time-inhomogeneous gamma density*.

*Proof.* This follows from the definition  $f(t) = -\bar{F}'(t)$  and by differentiating in (3).

An additional condition is required to ensure that the pdf in (4) and the survivor function in (3) define a proper probability distribution. It is immediate from (3) that  $\bar{F}(t)$  is the survivor function of a proper probability distribution if and only if  $\lim_{t \rightarrow \infty} [A(t)]^j / \exp\{A(t)\} = 0$  for  $j = 0, 1, \dots, n-1$ . Repeated application of L'Hôpital's rule shows that this condition is met if  $A(t) \rightarrow \infty$ , that is, if the integrated event rate grows unboundedly. If this condition is not satisfied—for example, if  $a(t)$  goes to zero in finite time—then the probability mass in the distribution in (4) will be less than 1. That is, there will be a non-zero probability that the process will never terminate, which will occur if the event rate goes to zero before the counting process  $N(t)$  reaches criterion.

### 3. The time-varying, two-alternative counter model

#### 3.1. Response time density functions

The assumed independence of the Poisson point processes  $X_A(t)$  and  $X_B(t)$  endows the parallel-counter model with a very simple probabilistic structure. For example, the pdf  $g_A(t)$  satisfies the relationship

$$\begin{aligned} g_A(t) dt &= P \left[ T_A \in (t - dt, t], T_B > t \right] \\ &= P \left[ T_A \in (t - dt, t] \right] P \left[ T_B > t \right], \end{aligned}$$

where the second equality follows by independence. The following proposition, which is one of the main results of this article, is then immediate from the definitions of  $T_A$  and  $T_B$  and of the time-varying gamma density and survivor functions in (4) and (3).<sup>3</sup>

*Proposition 3.1.* Let the integrated event rates for the counting processes associated with responses  $R_A$  and  $R_B$  be  $A(t)$  and  $B(t)$ , respectively. The joint RT density functions,  $g_A(t)$  and  $g_B(t)$ , of  $R_A$  and  $R_B$  are:

$$g_A(t) = \frac{[A(t)]^{K_A-1} A'(t) e^{-A(t)}}{(K_A - 1)!} \left\{ \sum_{j=0}^{K_B-1} \frac{[B(t)]^j}{j!} e^{-B(t)} \right\} \tag{5a}$$

and

$$g_B(t) = \frac{[B(t)]^{K_B-1} B'(t) e^{-B(t)}}{(K_B - 1)!} \left\{ \sum_{j=0}^{K_A-1} \frac{[A(t)]^j}{j!} e^{-A(t)} \right\}, \tag{5b}$$

respectively.

From these results we obtain the marginal RT density  $g(t)$ ,

$$g(t) = g_A(t) + g_B(t),$$

which is the pdf of  $T = \min(T_A, T_B)$ , the finishing time of the first-finishing process, and the response probabilities  $P(A)$  and  $P(B)$ ,

$$P(A) = \int_0^\infty g_A(t) dt$$

and

$$P(B) = \int_0^\infty g_B(t) dt,$$

which are the probabilities of responses  $R_A$  and  $R_B$ , respectively. The corresponding conditional RT density functions are obtained by dividing  $g_A(t)$  and  $g_B(t)$  by their associated response probabilities,  $P(A)$  and  $P(B)$ . When no restrictions are placed on the form of the Poisson rate functions  $a(t)$  and  $b(t)$ , closed-form expressions for these probabilities cannot be obtained. In most situations this represents only a minor inconvenience, since in the majority of applications of models of this kind, the RT distributions will be the quantities of primary

<sup>3</sup> The results in Proposition 3.1 were derived by each of the authors independently. Their first appearance, to our knowledge, was in Van Zandt's (1992) unpublished PhD thesis.

theoretical interest and, as noted previously, other statistics can be obtained from these by numerical integration. However, there is an important special case in which closed-form expressions for the choice probabilities are readily obtained.

*Definition.* We call a parallel-counter model a *proportional-rates model* if there exists a function  $U(t) \geq 0$  such that the integrated Poisson event rates  $A(t)$  and  $B(t)$  satisfy the relationship

$$\begin{aligned} A(t) &= a U(t), \\ B(t) &= b U(t), \end{aligned} \tag{6}$$

for positive, real constants,  $a$  and  $b$ . The standard (time-homogeneous) counter model studied by Townsend and Ashby (1983) satisfies this condition with  $U(t) = t$ .

Note that this condition, which is essentially a separability requirement, is a very natural one in many applications. In essence, it states that the Poisson rates  $a(t)$  and  $b(t)$  are a product of the quality of the (time-dependent) stimulus representation,  $U'(t)$ , and of the degree to which the stimulus matches the internal, memorial representations associated with each of the responses. In this, it closely resembles the notion of ‘resonance’ used by Ratcliff (1978) to motivate his diffusion process model of memory retrieval. It also resembles the ‘proportional hazards’ assumption that is used widely in models in survival analysis (Cox & Oakes, 1984).

### 3.2. Response probabilities

The response probabilities  $P(A)$  and  $P(B)$  are obtained by integrating the pdfs in (5a) and (5b) from zero to infinity. For example,

$$P(A) = \sum_{j=0}^{K_B-1} \int_0^{\infty} \frac{[A(s)]^{K_A-1} A'(s) e^{-A(s)}}{(K_A-1)!} \left\{ \frac{[B(s)]^j}{j!} e^{-B(s)} \right\} ds,$$

which becomes, under the proportional-rates assumption,

$$\begin{aligned} P(A) &= \sum_{j=0}^{K_B-1} a^{K_A} b^j \int_0^{\infty} \frac{[U(s)]^{K_A+j-1} U'(s) e^{-(a+b)U(s)}}{(K_A-1)! j!} ds \\ &= \sum_{j=0}^{K_B-1} \frac{(K_A+j-1)!}{(K_A-1)! j!} \frac{a^{K_A} b^j}{(a+b)^{K_A+j}} \\ &\quad \times \int_0^{\infty} \frac{[(a+b)U(s)]^{K_A+j-1} (a+b)U'(s) e^{-(a+b)U(s)}}{(K_A+j-1)!} ds. \end{aligned}$$

The change of variable  $z = (a+b)U(s)$  reduces the integrand on the right-hand side of the last expression to a standard  $(K_A+j)$ -stage gamma pdf, which, under the condition  $\lim_{s \rightarrow \infty} U(s) = \infty$  of Lemma 2.1, integrates to unity. The probability of response  $R_A$  is therefore simply

$$P(A) = \sum_{j=0}^{K_B-1} \binom{K_A+j-1}{j} \left( \frac{a}{a+b} \right)^{K_A} \left( \frac{b}{a+b} \right)^j.$$

This expression is identical to that derived for the standard, time-homogeneous Poisson counter model by Townsend & Ashby (1983, p. 275). This equivalence highlights an

interesting property of the proportional-rates model, namely, that when the integrated event rate is unbounded, the response probabilities depend only on the relative magnitudes of the Poisson rate constants  $a$  and  $b$  and not on the form of the temporal dependency,  $U(t)$ .

The response probabilities for the proportional-rates case can be expressed in a more succinct form, by recalling the form of the negative binomial distribution, which we denote here by  $b^*(n, m, \pi)$ :

$$b^*(n, m, \pi) = \binom{n-1}{m-1} \pi^m (1-\pi)^{n-m}.$$

The negative binomial distribution gives the probability of waiting  $n$  Bernoulli trials for the occurrence of  $m$  ‘successes’, when the probability of a success in a single trial is  $\pi$ . The response probabilities for the proportional-rates model may be expressed as a sum of negative binomial terms. We record the result in the form of a proposition.

*Proposition 3.2.* Let the Poisson rate constants associated with responses  $R_A$  and  $R_B$  be  $a$  and  $b$ , respectively. Define

$$\begin{aligned} \pi_a &= \frac{a}{a+b}, \\ \pi_b &= \frac{b}{a+b}, \end{aligned}$$

and assume that  $\lim_{t \rightarrow \infty} U(t) = \infty$ . Then the response probabilities for the proportional-rates model are

$$P(A) = \sum_{j=0}^{K_B-1} b^*(K_A + j, K_A, \pi_a) \tag{7a}$$

and

$$P(B) = \sum_{j=0}^{K_A-1} b^*(K_B + j, K_B, \pi_b). \tag{7b}$$

As pointed out by LaBerge (1962, 1994), an alternative, integral representation of the response probabilities may also be obtained (see also Luce, 1986, pp. 320–322). This is based on the well-known relationship between the cumulative of the negative binomial and the incomplete beta function. We recall the definition of the incomplete beta function,  $I_\pi(n, m)$  (Gradshteyn & Ryzhik, 1965, eqs. 8.390–8.392). For  $n$  and  $m$  integer,  $0 \leq \pi \leq 1$ :

$$I_\pi(n, m) = n \binom{n+m-1}{m-1} \int_0^\pi z^{n-1} (1-z)^{m-1} dz.$$

It may then be shown that  $P(A)$  and  $P(B)$  in (7a) and (7b) have the integral representations

$$P(A) = I_{\pi_a}(K_A, K_B) \tag{8a}$$

and

$$P(B) = I_{\pi_b}(K_B, K_A). \tag{8b}$$

With these representations, a simple relationship between the Poisson rate parameters and the ordering of response probabilities can readily be established. We refer to a model as *unbiased* if the response criteria in the evidence counters are equal, that is, if  $K_A = K_B$ .

*Proposition 3.3.* For all unbiased proportional-rates models in which the integrated event rate grows unboundedly,  $P(A) \geq P(B)$  if  $a \geq b$ .

*Proof.* When the integrated event rate is unbounded (i.e.,  $T$  is finite with probability 1), the response probabilities are given by the incomplete beta representations (8a) and (8b) with  $K_A = K_B$ . Because the integrand in the incomplete beta function is positive, the integrand is increasing in its upper bound of integration  $\pi$ . Therefore  $a \geq b$ , which implies  $\pi_a \geq \pi_b$ , implies  $P(A) \geq P(B)$ .

### 3.3. Mean response times

We denote by  $E[T|R_A]$  and  $E[T|R_B]$  the mean RTs for responses  $R_A$  and  $R_B$ , respectively. Using a standard property of survivor functions, these may be calculated as

$$E[T|R_A] = \frac{1}{P(A)} \int_0^{\infty} \bar{G}_A(t) dt \quad (9a)$$

$$E[T|R_B] = \frac{1}{P(B)} \int_0^{\infty} \bar{G}_B(t) dt, \quad (9b)$$

where  $\bar{G}_A(t)$  and  $\bar{G}_B(t)$  are, respectively, the survivor functions associated with the joint pdfs  $g_A(t)$  and  $g_B(t)$ . We consider the properties of the function  $\bar{G}_A(t)$ , the function for the other response being obtained in an analogous way.

From equation (5a),

$$\bar{G}_A(t) = \sum_{j=0}^{K_B-1} \int_t^{\infty} \frac{A'(s)[A(s)]^{K_A-1}[B(s)]^j e^{-[A(s)+B(s)]}}{(K_A-1)!j!} ds.$$

For the proportional-rates case,

$$\begin{aligned} \bar{G}_A(t) &= \sum_{j=0}^{K_B-1} \int_t^{\infty} \frac{aU'(s)[aU(s)]^{K_A-1}[bU(s)]^j e^{-(a+b)U(s)}}{(K_A-1)!j!} ds \\ &= \sum_{j=0}^{K_B-1} \frac{(K_A+j-1)!}{(K_A-1)!j!} \frac{a^{K_A} b^j}{(a+b)^{K_A+j}} \\ &\quad \times \int_t^{\infty} \frac{(a+b)U'(s)[(a+b)U(s)]^{K_A+j-1}}{(K_A+j-1)!} e^{-(a+b)U(s)} ds. \end{aligned}$$

To express this function more compactly, we recall the definition of the gamma function (Gradshteyn & Ryzhik, 1965, eqs. 8.310, 8.238),

$$\begin{aligned} \Gamma(p) &= \int_0^{\infty} z^{p-1} e^{-z} dz \\ &= (p-1)! \quad \text{for } p = 1, 2, \dots; \end{aligned}$$

and the incomplete gamma function (Gradshteyn & Ryzhik, 1965, eqs. 8.350, 8.352),

$$\begin{aligned} \Gamma(p, t) &= \int_t^\infty z^{p-1} e^{-z} dz \\ &= \Gamma(p) \sum_{j=0}^{p-1} \frac{t^j}{j!} e^{-t} \quad \text{for } p = 1, 2, \dots \end{aligned} \quad (10)$$

By making the change of variable  $z = (a + b)U(s)$  and recalling the previous notation for the negative binomial distribution, the survivor function  $\tilde{G}_A(t)$  may be expressed in the compact form

$$\tilde{G}_A(t) = \sum_{j=0}^{K_B-1} b^*(K_A + j, K_A, \pi_a) \frac{\Gamma[K_A + j, (a + b)U(t)]}{\Gamma(K_A + j)}.$$

From this expression and its counterpart for  $\tilde{G}_B(t)$ , a general representation for the conditional mean RTs for the counter model may be obtained by substitution in (9a) and (9b), which we again record in the form of a proposition.

*Proposition 3.4.* The conditional mean RTs for the proportional-rates Poisson counter model are of the form

$$E[T|R_A] = \frac{1}{P(A)} \sum_{j=0}^{K_B-1} \frac{b^*(K_A + j, K_A, \pi_a)}{\Gamma(K_A + j)} \int_0^\infty \Gamma[K_A + j, (a + b)U(s)] ds, \quad (11a)$$

and

$$E[T|R_B] = \frac{1}{P(B)} \sum_{j=0}^{K_A-1} \frac{b^*(K_B + j, K_B, \pi_b)}{\Gamma(K_B + j)} \int_0^\infty \Gamma[K_B + j, (a + b)U(s)] ds. \quad (11b)$$

Although these equations give a general representation for the mean RTs, in the majority of cases, no closed-form expression form will exist for the integral with respect to time of the incomplete gamma as a function of a time-varying parameter. There is, however, one interesting special case for which closed-form expressions may easily be derived. This is the case in which the integrated event rate function is a power function of the form  $U(t) = t^c$ , for some real constant  $c$ ,  $c > 0$ .

Making use of the definition of the incomplete gamma function (10), the expression for  $E[T|R_A]$  may be written in explicit form as

$$\begin{aligned} E[T|R_A] &= \frac{1}{P(A)} \sum_{j=0}^{K_B-1} b^*(K_A + j, K_A, \pi_a) \\ &\quad \times \sum_{k=0}^{K_A+j-1} \int_0^\infty \frac{[(a + b)U(s)]^k}{k!} e^{-(a+b)U(s)} ds. \end{aligned}$$

When  $U(t)$  is a power function, the integral term on the right of the preceding equation, which we denote  $Z(c, k)$ , is of the form

$$Z(c, k) = \int_0^\infty \frac{(a + b)^k t^{ck}}{k!} \exp[-(a + b)t^c] dt.$$

From the standard integral (Gradshteyn & Ryzhik, 1965, eq. 3.478)

$$\int_0^{\infty} x^{v-1} \exp(-\mu x^p) dx = \frac{1}{p} \mu^{-v/p} \Gamma(v/p),$$

we have

$$Z(c, k) = \frac{1}{c} (a + b)^{-1/c} \frac{\Gamma(k + 1/c)}{\Gamma(k + 1)}.$$

Substituting this in the expression for  $E[T|R_A]$  and proceeding analogously with  $E[T|R_B]$  we obtain the following proposition.

*Proposition 3.5.* When the integrated event rate is a power function, the conditional mean RTs for the proportional-rates model are

$$E[T|R_A] = \frac{1}{P(A)} \sum_{j=0}^{K_B-1} b^*(K_A + j, K_A, \pi_a) \sum_{k=0}^{K_A+j-1} \frac{\Gamma(k + 1/c)}{\Gamma(k + 1)} \frac{1}{c} (a + b)^{-1/c}, \quad (12a)$$

and

$$E[T|R_B] = \frac{1}{P(B)} \sum_{j=0}^{K_A-1} b^*(K_B + j, K_B, \pi_b) \sum_{k=0}^{K_B+j-1} \frac{\Gamma(k + 1/c)}{\Gamma(k + 1)} \frac{1}{c} (a + b)^{-1/c}. \quad (12b)$$

In the special case of time-homogeneous accrual processes,  $c = 1$  and  $\Gamma(k + 1/c)/\Gamma(k + 1) = 1$ . Then the summations of terms over  $k$  in the preceding expressions reduce to  $(K_A + j)/(a + b)$  and  $(K_B + j)/(a + b)$ , respectively, and the equations become identical to those derived for the homogeneous model by Townsend & Ashby (1983, p. 276).

The next result concerns the ordering of mean RTs for correct responses and errors. It shows that when the response criteria for the counters in a proportional-rates model are equal, an ordering of mean RTs is induced by the ordering of Poisson rate constants.

*Proposition 3.6.* For all unbiased proportional-rates models,  $E[T|R_A] \leq E[T|R_B]$  if  $a \geq b$ .

To prove this proposition we need a definition and a lemma.

*Definition.* For a pair of discrete random variables  $X$  and  $Y$ ,  $X$  is said to be stochastically greater than  $Y$ , written  $X \geq_{\text{st}} Y$ , if  $P[X \geq j] \geq P[Y \geq j]$  for all  $j$  in the domain of  $X$  and  $Y$ . That is,  $X \geq_{\text{st}} Y$  means the distribution function of  $Y$  dominates that of  $X$ .

*Lemma 3.1.* For a pair of random variables  $X$  and  $Y$ ,  $X \geq_{\text{st}} Y$  if and only if  $E[f(X)] \geq E[f(Y)]$  for all increasing functions  $f$ .

This result is proved in Ross (1983, Proposition 8.1.2).

*Proof of Proposition 3.6.* Let  $K_A = K_B = K$ . With this constraint, the conditional mean RTs in (11a) and (11b) may be written

$$\begin{aligned} E[T|R.] &= E \left\{ \frac{\int_0^{\infty} \Gamma[K + J(\pi.), (a + b)U(s)] ds}{\Gamma[K + J(\pi.)]} \right\} \\ &= \sum_{j=0}^{K-1} P[J(\pi.) = j] \frac{\int_0^{\infty} \Gamma[K + j, (a + b)U(s)] ds}{\Gamma[K + j]}. \end{aligned}$$

This expectation is of the form  $E[f\{J(\pi.)\}]$ , where  $J(\pi.)$ , is a conditional negative binomial

random variable with probability law

$$P[J(\pi) = j] = \frac{b^*(K + j, K, \pi)}{\sum_{k=0}^{K-1} b^*(K + k, K, \pi)}; \quad j = 0, 1, \dots, K - 1. \quad (13)$$

By assumption,  $(a + b)U(t)$  is positive, and we note from the definition of the incomplete gamma function (10) that under these circumstances, for integer  $K + j$ , the function

$$f(j) = \frac{\int_0^\infty \Gamma[K + j, (a + b)U(s)] ds}{\Gamma[K + j]} \quad (14)$$

is a sum of positive terms that is increasing in  $j$ . By Lemma 3.1,  $E[T|R_A] \leq [E T|R_B]$  if  $J(\pi_b) \geq_{st} J(\pi_a)$ .

To show that  $J(\pi_b) \geq_{st} J(\pi_a)$  we make use of the one-point crossover property described by Townsend (1990) and Townsend & Ashby (1983, pp. 281–282). For a pair of discrete random variables  $X$  and  $Y$ , a sufficient condition for  $Y \geq_{st} X$  is that (a)  $P[X = j'] \geq P[Y = j']$  for the smallest index  $j'$  for which  $P[Y = j] \neq 0$ , and (b) the function  $P[X = j] - P[Y = j]$  changes sign at most once on the domain  $j > j'$ .<sup>4</sup> From the definition of  $P[J(\pi)]$  in (13),

$$P[J(\pi_a) = j] - P[J(\pi_b) = j] = \frac{b^*(K + j, K, \pi_a)}{\sum_{k=0}^{K-1} b^*(K + k, K, \pi_a)} - \frac{b^*(K + j, K, \pi_b)}{\sum_{k=0}^{K-1} b^*(K + k, K, \pi_b)}.$$

At  $j = 0$  this reduces to

$$\begin{aligned} & P[J(\pi_a) = 0] - P[J(\pi_b) = 0] \\ &= \pi_a^K \left[ \sum_{k=0}^{K-1} \binom{K+k-1}{k} \pi_a^K (1 - \pi_a)^k \right]^{-1} - \pi_b^K \left[ \sum_{k=0}^{K-1} \binom{K+k-1}{k} \pi_b^K (1 - \pi_b)^k \right]^{-1} \\ &= \left[ \sum_{k=0}^{K-1} \binom{K+k-1}{k} \pi_b^k \right]^{-1} - \left[ \sum_{k=0}^{K-1} \binom{K+k-1}{k} \pi_a^k \right]^{-1}, \end{aligned}$$

by virtue of the fact that  $\pi_a + \pi_b = 1$ . This last expression is of the form

$$P[J(\pi_a) = 0] - P[J(\pi_b) = 0] = \left[ \sum_{k=0}^{K-1} w_k \pi_b^k \right]^{-1} - \left[ \sum_{k=0}^{K-1} w_k \pi_a^k \right]^{-1},$$

where the  $w_k$  are positive constants. Since, by assumption,  $\pi_b < \pi_a$  we have  $(\sum w_k \pi_b^k)^{-1} > (\sum w_k \pi_a^k)^{-1}$ , which means that  $P[J(\pi_a) = 0] - P[J(\pi_b) = 0] > 0$ , which verifies part (a). For part (b) we note that  $P[J(\pi_a) = j] - P[J(\pi_b) = j]$  is proportional to

$$\binom{K+j-1}{j} \left[ \pi_a^K (1 - \pi_a)^j - c \pi_b^K (1 - \pi_b)^j \right]; \quad j = 0, 1, \dots, K - 1, \quad (15)$$

<sup>4</sup> The stochastic dominance results in Townsend (1990) and Townsend & Ashby (1983) were derived for pairs of continuous random variables and were expressed in terms of the crossover properties of their pdfs. However, these same properties apply with no essential modifications to the probability mass functions of discrete random variables also.

where

$$c = \frac{\sum_{k=0}^{K-1} b^*(K+k, K, \pi_a)}{\sum_{k=0}^{K-1} b^*(K+k, K, \pi_b)}$$

is a positive constant that is independent of  $j$ . Since  $(1 - \pi_a) < (1 - \pi_b) < 1$ , the quantity in square brackets in (15) is strictly decreasing in  $j$ . Therefore  $P[J(\pi_a) = j] - P[J(\pi_b) = j]$  changes sign at most once on  $j \geq 0$ , from which the result  $J(\pi_b) \geq_{\text{st}} J(\pi_a)$  follows.

### 3.4. The $m$ -alternative case

As a result of its simple, independent channels structure, the model extends in a straightforward way to the  $m$ -alternative case. Here we give a general expression for the joint RT density functions and the response probabilities and mean RTs for the proportional-rates case. To this end, we modify the previous notation slightly and denote the channel rates by  $a_i(t)$  and the integrated channel rates by  $A_i(t)$ ,  $i = 1, 2, \dots, m$ . As in previous sections, in the proportional-rates case the channel rates are  $a_i U'(t)$  for some function  $U(t)$ . For this case, we denote by  $a_\Sigma$  the sum of the rate constants  $\sum_i a_i$ . We denote the joint RT density function for response  $R_i$  by  $g_i(t)$  and its associated response probability by  $P(R_i)$ .

By a straightforward generalization of the argument leading to (5a) and (5b) we obtain the following result.

*Proposition 3.7.* The joint RT density function for response  $R_i$  for the general,  $m$ -alternative, Poisson counter model is

$$g_i(t) = \frac{[A_i(t)]^{K_i-1} A_i'(t) e^{-A_i(t)}}{(K_i - 1)!} \prod_{\substack{j=1 \\ j \neq i}}^m \left\{ \sum_{k=0}^{K_j-1} \frac{[A_j(t)]^k}{k!} e^{-A_j(t)} \right\}. \quad (16)$$

The response probability  $P(R_i)$  is obtained by integrating this density over the positive real line. For the proportional-rates model, the expression is

$$\begin{aligned} P(R_i) &= \int_0^\infty \frac{[a_i U(t)]^{K_i-1} a_i U'(t) e^{-a_i U(t)}}{(K_i - 1)!} \prod_{\substack{j=1 \\ j \neq i}}^m \left\{ \sum_{k=0}^{K_j-1} \frac{[a_j U(t)]^k}{k!} e^{-a_j U(t)} \right\} dt \\ &= \int_0^\infty \frac{[a_i U(t)]^{K_i-1} a_i U'(t) e^{-a_\Sigma U(t)}}{(K_i - 1)!} \\ &\quad \times \left\{ \underbrace{\sum_{k_m=0}^{K_m-1} \dots \sum_{k_1=0}^{K_1-1} \frac{[a_i U(t)]^{K_i-1} [a_1 U(t)]^{k_1} \dots [a_m U(t)]^{k_m}}{(K_i - 1)! k_1! \dots k_m!}}_{m-1 \text{ terms}} \right\} dt \\ &= \sum_{k_m=0}^{K_m-1} \dots \sum_{k_1=0}^{K_1-1} \frac{(K_i + k_1 + \dots + k_m - 1)!}{(K_i - 1)! k_1! \dots k_m!} \frac{a_i^{K_i} a_1^{k_1} \dots a_m^{k_m}}{(a_\Sigma)^{K_i+k_1+\dots+k_m}} \\ &\quad \times \int_0^\infty \frac{(a_\Sigma)^{K_i+k_1+\dots+k_m} U'(t) e^{-a_\Sigma U(t)} U(t)^{K_i+k_1+\dots+k_m-1}}{(K_i + k_1 + \dots + k_m - 1)!} dt. \end{aligned}$$

In the preceding expressions, occurrences of the ellipsis “...” are to be understood as denoting iterated summation over a set of  $m$  subscripts  $k_1, k_2, \dots, k_m$ , excluding the  $i$ th subscript.

By making the change of variable,  $z = a_\Sigma U(t)$ , the integral on the right-hand side of the last equality may be reduced to that of a standard  $(K_i + k_1 + \dots + k_m)$ -stage gamma density. By assumption,  $U(0) = 0$ , and if the integrated event rate grows unboundedly,  $\lim_{t \rightarrow \infty} U(t) = \infty$ , the integral will be unity. Under these circumstances, to express the result more compactly, we define, by analogy with the negative binomial, the *negative multinomial distribution*, which we denote by  $m^*(n_1, \dots, n_m; r; \pi_1, \dots, \pi_m)$ . This distribution gives the probability of waiting a total of  $\sum_i n_i$   $m$ -valued trials for  $r$  successes, where a ‘success’ refers to the occurrence of a trial of one particular type. We define  $\pi_i = a_i/a_\Sigma$ ,  $i = 1, 2, \dots$ . With this notation, the probability of response  $R_i$  may be written as a sum of negative multinomial probabilities where the summation is taken over all possible sequences of event types that lead to response  $R_i$ .

*Proposition 3.8.* For the  $m$ -alternative proportional-rates case, when the integrated event rate is unbounded, the proportion of  $R_i$  responses is

$$P(R_i) = \underbrace{\sum_{k_m=0}^{K_m-1} \dots \sum_{k_1=0}^{K_1-1}}_{m-1 \text{ terms}} m^*(k_1, \dots, K_i, \dots, k_m; K_i; \pi_1, \dots, \pi_m). \tag{17}$$

As in the preceding expressions, the iterated summation is understood to exclude the  $i$ th subscript.

When the integrated event rate is a power function, the conditional mean RTs  $E[T|R_i]$  may be derived by a method that parallels that for the two-alternative case. This involves a double integration over the density function, once to obtain the survivor function and once to obtain the expected value. We omit the details and merely record the result.

*Proposition 3.9.* For the  $m$ -alternative proportional-rates case, when the integrated event rate is of the form  $U(t) = t^c$ , the conditional mean RT for response  $R_i$  is

$$E[T|R_i] = \frac{1}{P(R_i)} \underbrace{\sum_{k_m=0}^{K_m-1} \dots \sum_{k_1=0}^{K_1-1}}_{m-1 \text{ terms}} m^*(k_1, \dots, K_i, \dots, k_m; K_i; \pi_1, \dots, \pi_m) \\ \times \sum_{j=0}^{K_i+k_1+\dots+k_m-1} \frac{\Gamma(j + 1/c)}{\Gamma(j + 1)} \frac{1}{c} (a_\Sigma)^{-1/c}. \tag{18}$$

### 4. Empirical application

#### 4.1. Sustained and transient channels

As noted previously, one area in which models with time-varying accumulation rates arise naturally is in visual psychophysics. Models of early visual processing typically assume that the coding of elementary stimulus attributes in the visual system is carried out by an array of retinotopically distributed cortical filters, each of which is tuned for spatial frequency and

orientation. These filters are often represented mathematically, at least to a first approximation, as simple linear systems. The response of such a filter to an arbitrary temporal waveform is given by the convolution of the waveform with the impulse response function of the filter. In general, this response will not be constant, but will vary continuously with time. There is a large literature that relates both psychophysical and physiological data to physical stimulus attributes using models of this kind, a good introduction to which may be found in Watson (1986).

A fundamental distinction that emerges from this literature is between *sustained* and *transient* processing channels. This distinction, which has been investigated experimentally by numerous authors (e.g., Legge, 1978; Tolhurst, 1975a) refers to systems that respond, on the one hand, to the steady-state luminance properties of a stimulus and, on the other, to the temporal change in luminance that occurs at stimulus onset and offset. Evidence for the existence of these two classes of mechanism comes from a variety of sources, including suprathreshold RT (Tolhurst, 1975b), appearance thresholds (Kulikowski & Tolhurst, 1973), pulse-pair summation (Watson & Nachmias, 1977) and estimates of the visual modulation transfer function in flicker-detection experiments (Robson, 1966). Physiologically, the sustained–transient dichotomy appears to be an expression, at least in part, of the anatomical and functional division of the primate visual system into parvocellular (P) and magnocellular (M) pathways. Evidence linking sustained and transient behaviour to the properties of P and M systems comes from single-cell recording (Purpura, Kaplan & Shapley, 1988; Valberg & Lee, 1989) and pharmacological lesion studies (Merigan, Byrne & Maunsell, 1991).

In the literature, two distinct approaches to modelling the properties of sustained and transient channels may be found, one of which treats the degree of transience of the response as a continuum (e.g., Busey, 1999; Watson, 1986) and another that treats it as essentially dichotomous (e.g., Fredericksen & Hess, 1997; Gorea & Tyler, 1986; Smith, 1995). The justification for treating it as a continuum is the empirical finding that single cells in both the P and M systems show large individual differences in the degree of transience of their responses. The justification for treating it as dichotomous is that the psychophysical literature on visual temporal sensitivity provides little evidence for more than two qualitatively distinct classes of mechanism. Although the mathematical methods described in this paper are compatible with representations of either kind, a dichotomous structure is assumed here, both because of its conceptual simplicity and because it allows the present results to be compared directly with those of previous efforts to link the properties of sustained and transient mechanisms to sequential-sampling decision models in a diffusion process setting (Smith, 1995, 1998). The details are as follows.

The luminance (or contrast) of an arbitrary stimulus may be represented as a function  $l(t, x, y)$  of time and of its horizontal and vertical coordinates in the visual field. For simplicity, however, the dependence on spatial position will be omitted from the notation and the function denoted by  $l(t)$ . Since de Lange (1954, 1958), research on temporal sensitivity has typically assumed that the response to an arbitrary, space-time separable visual stimulus may be modelled as the output of a time-invariant, causal, linear system. The sustained system is usually modelled as a set of low-pass filters, which block or attenuate all temporal frequencies that exceed a specified cutoff value. Mathematically, such a filter may be represented as a linear operator,  $L_S$ , whose response to stimulus  $l(t)$  is denoted  $\mu_S(t)$ , where the subscript in either instance identifies the sustained system. The relationship between the input and output

of the filter may be written in operator notation as

$$\mu_S(t) = L_S l(t).$$

Because  $L_S$  is a low-pass filter, asymptotically its output in response to constant input will be proportional to  $l(t)$ . Moreover, an abrupt change in  $l(t)$  will produce a corresponding change in output, but at a slower rate than the change in  $l(t)$  because of the attenuation of the high temporal frequency components of  $l(t)$  by the filter.

The transient system, in contrast, is usually modelled as a set of band-pass filters, which block or attenuate frequencies that are above or below the upper and lower cutoffs of the filter pass band. One way to realize such a filter is to combine a differentiator,  $D_T$ , and a low-pass filter  $L_T$ , in cascade (Smith, 1995, 1998), where the subscript in either instance identifies the transient system. The output of the composite system, denoted  $\mu_T(t)$ , may be written

$$\mu_T(t) = L_T D_T l(t) = L_T l'(t).$$

The differentiation stage in the preceding equation means that the transient system responds only to changes in the value of  $l(t)$ , because the output of a differentiator to constant input is zero. The presence of a low-pass filtering stage means that, like the sustained system, the temporal resolution of the filter in response to such changes is limited. In psychophysical applications, however, it is typically assumed that the degree of high frequency attenuation of  $L_T$  is appreciably less than that of  $L_S$ . This endows the transient system with a relatively brisk temporal response in comparison to that of the sustained system, a property which accords with its assumed biological function as a system specialized for the detection of rapid environmental changes.

Mathematically, the systems  $L_S$  and  $D_T L_T$  may be characterized uniquely via their temporal impulse response functions, which describe their outputs to a Dirac delta input,  $\delta(t)$ . Equivalently, they may be characterized by their responses to a Heaviside unit-step function  $u(t)$ , where the latter is defined as

$$u(t) = \begin{cases} 1, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

The resultant outputs, which are called the *step-response functions* of the sustained and transient systems, will be denoted by  $\nu_S(t)$  and  $\nu_T(t)$ , respectively.

In the psychophysical modelling literature, the low-pass filtering properties of early visual mechanisms are often represented by a set of cascaded R-C (exponential) stages, that is, by a sequence of processing stages, each of which is governed by a first-order, linear differential equation (e.g., Smith, 1995, 1998; Sperling & Sondhi, 1968; Watson, 1986). Any filter of this kind is characterized by three parameters,  $n$ ,  $\alpha$  and  $\beta$ , which specify, respectively, the number of cascaded stages comprising the filter, its gain or amplitude response, and the common stage rate constant. These parameter pairs will be denoted  $(n_S, \alpha_S, \beta_S)$  for the sustained system and  $(n_T, \alpha_T, \beta_T)$  for the transient system. With the assumptions given above, the step-response function of the sustained system is

$$\nu_S(t) = L_S u(t) = \alpha_S \left[ 1 - e^{\beta_S t} \sum_{j=0}^{n_S-1} \frac{(\beta_S t)^j}{j!} \right] \quad (19)$$

and that of the transient system is

$$\nu_T(t) = D_T L_T u(t) = \alpha_T \frac{(\beta_T t)^{n_T-1}}{(n_T - 1)!} \beta_T e^{-\beta_T t}. \quad (20)$$

The qualitative behaviour of the functions in (19) and (20) accords with the properties usually ascribed to these systems for abrupt-onset, response-terminated stimuli. The sustained response,  $\nu_S(t)$ , increases smoothly to an asymptote, while the transient response,  $\nu_T(t)$ , increases to a peak and then decays back to zero. The parameters  $\alpha_S$  and  $\alpha_T$  are assumed to depend jointly on signal strength and the degree to which the spatial configuration of the stimulus matches the receptive field of the associated filter. Note that the functions  $\nu_S(t)$  and  $\nu_T(t)$  are identical in form, up to a scaling constant, to the distribution function and the density function of the gamma probability distribution, discussed (in generalized form) in preceding sections. Here, however, rather than representing the probabilistic characteristics of a sum of exponentially distributed random variables, they represent the continuous, deterministic output of early visual filters.

The full specification of a model of perceptual decision-making requires a further, auxiliary, hypothesis that links the (deterministic) filter output with the stochastic properties of the decision stage. A natural way to do this is to assume that the neural spike densities in the channels that encode the stimulus are proportional to the filter output. In the case of the proportional-rates model, this is tantamount to assuming that

$$U(t) = \int_0^t \mu_S(s) ds \quad \text{or} \quad U(t) = \int_0^t \mu_T(s) ds$$

according to whether the filter in question is sustained or transient, with the filter amplitudes  $\alpha_S$  or  $\alpha_T$  set to  $a$  and  $b$  for the  $R_A$  and  $R_B$  channels, respectively. For abrupt-onset stimuli presented at time  $t = 0$  and terminated by a response,  $\mu_S(t)$  and  $\mu_T(t)$  are adequately approximated by the functions  $\nu_S(t)$  and  $\nu_T(t)$  in (19) and (20).

Figures 1(a) and 1(b) show RT density functions for a parallel Poisson model driven by sustained and transient channels using stimuli of this kind. The functions shown are the joint density functions  $g_A(t)$  and  $g_B(t)$ , where, in all cases, the correct (i.e., more frequent) response is  $R_A$ . All of the density functions shown are for an unbiased decision process,  $K_A = K_B$ . Figures 1(c) and 1(d) show the corresponding hazard functions, which are defined as

$$h_i(t) = \frac{g_i(t)}{G_i(t)}; \quad i = A, B.$$

In all panels of the figure, the parameter of the family of functions shown is the channel rate parameter ( $\beta_S$  or  $\beta_T$ ). These parameters determine how rapidly the channel outputs reach their maximum once the stimulus is presented and, in the case of the transient channel, how rapidly the channel output then decays to zero. In agreement with the results of Section 3.3, for all parameter combinations shown, mean correct RTs are faster than error RTs, a property which the proportional-rates model shares with the time-homogeneous model studied by Townsend & Ashby (1983).

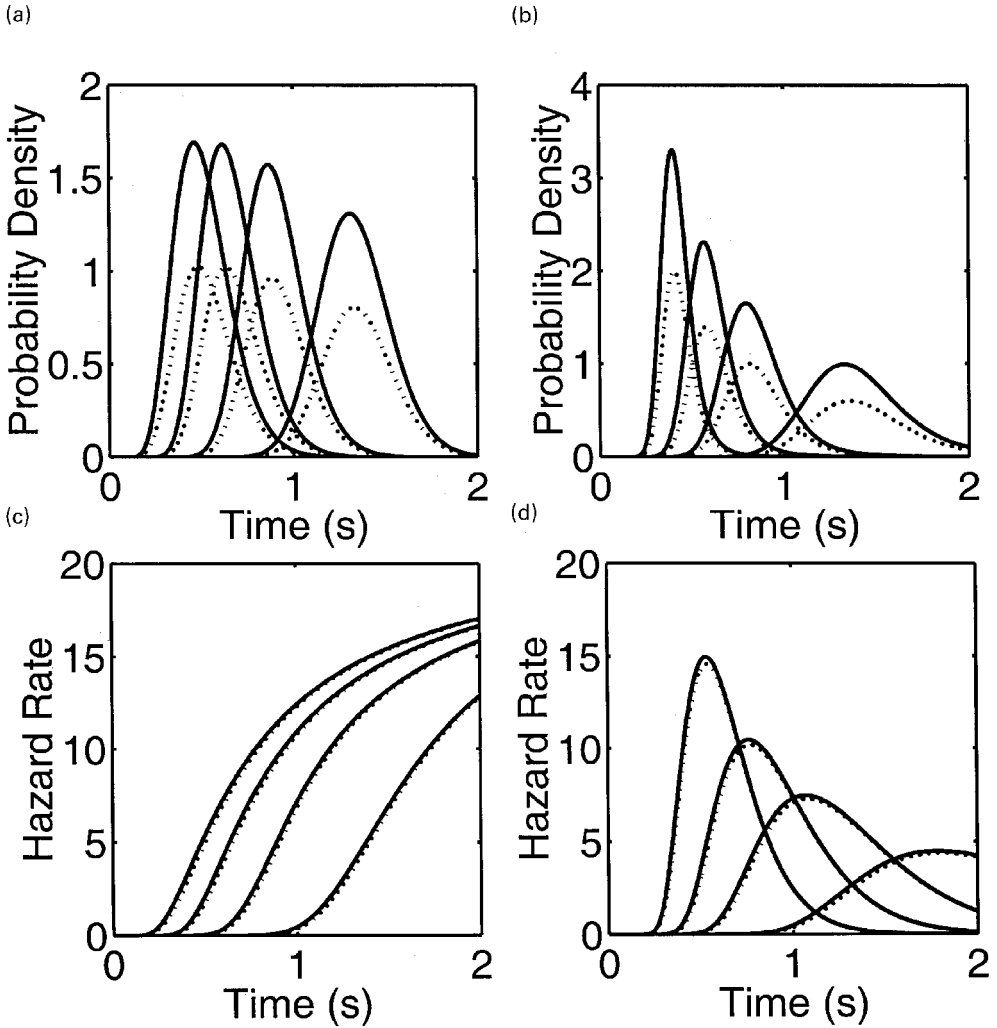
Another important property of the RT distributions in Fig. 1(a) is that, for large values of the channel rate parameter  $\beta_S$ , the RT distributions are related approximately by a simple shifting on the time axis. This property arises because the integrated event rate  $U(t)$  in (5a) and (5b) is given by the integral of  $\nu_S(t)$  in (19), which is asymptotically linear with slope  $\alpha_S$ .

Indeed, the asymptotic portion of the function is well described by a straight line of the form  $e(t) = \alpha_S(t - t_0)$ , where the shift  $t_0$  is jointly determined by the number of cascaded stages in the channel  $n_S$  and the channel rate constant  $\beta_S$ . For constant  $n_S$ , the larger the value of  $\beta_S$ , the better the linear approximation and the smaller the shift constant  $t_0$ . Under these circumstances, the model behaves asymptotically like the homogeneous model studied by Townsend & Ashby (1983) and the RT distributions are well approximated by the simpler expressions that apply in the homogeneous case, but shifted to the right by an amount  $t_0$ . This justifies the usual procedure of fitting the simpler, homogeneous model with an offset parameter  $t_0$  to data from experimental paradigms in which the task-relevant information is carried primarily by the sustained (i.e., high spatial frequency) components of the stimulus. This would be expected to apply, for example, in character classification tasks, which was the context in which Townsend & Ashby considered the model originally.

The hazard functions in Fig. 1(c) provide another perspective on the distributional characteristics of models driven by sustained channels. These show that when the integrated event rate is unbounded, the hazard functions are increasing, a property they share with those of the homogeneous model. This property is a straightforward consequence of the fact that, when  $a$  and  $b$  are non-zero and  $U(t) \rightarrow \infty$ , the evidence counters  $N_A(t)$  and  $N_B(t)$  fill progressively with the passage of time. The fuller the counters, the more likely it is that either  $K_A$  or  $K_B$  will be exceeded in the next instant and the process will terminate. The quantity  $h_i(t) dt$  approximates the probability that the process will terminate in the interval  $(t, t + dt]$ , given that it has not terminated by time  $t$ . Manifestly, this probability will increase as the counters fill, resulting in an increasing hazard function for both correct responses and errors.

The RT density functions for the transient system in Fig. 1(b) and the associated hazard functions in Fig. 1(d) give a rather different picture. As the channel rate constant  $\beta_T$  is increased, RTs become faster, as indicated by a shift of the density functions to the left along the time axis. However, the system as a whole also becomes more transient, as evidenced by an associated reduction in variance. The increased transience also appears clearly in the hazard functions in Fig. 1(d). Unlike the hazard functions for sustained channels, those for transient channels are non-monotonic and unimodal, increasing to a peak and then decaying back to zero. The unimodality of the hazard function means that there is some finite  $t$  at which the process is maximally likely to terminate. This is the time at which the system is most likely to register the occurrence of the stimulus onset transient. If no response is made at this time, the probability of a response being made subsequently will decrease as the time elapsing since the onset of the stimulus increases.

The properties of the transient system are such that, in general,  $P(A) + P(B) < 1$ , which means there will be a proportion of trials on which no response is made. However, in empirical applications of models of this kind, it is typically assumed that the transient system does not operate in isolation, but acts in concert with the sustained system, either in parallel, or through channel pooling. Multichannel, sequential-sampling models of this kind, in which RT depends on the simultaneous action of sustained and transient systems, have been investigated in a diffusion process setting by Smith (1995, 1998, in press), elaborating an idea proposed by Burbeck & Luce (1982). In these models, it is typically assumed that the processing rate of the transient system is faster than that of the sustained system, an assumption that is justified empirically by the faster neural conduction times in the primate *M* pathway. In composite models of this nature, the fastest RTs are mediated by transient



**Figure 1.** Response time density and hazard functions for a Poisson counter model driven by sustained and transient channels. (a) Density functions, sustained channel. (b) Density functions, transient channel. (c) Hazard functions, sustained channel. (d) Hazard functions, transient channel. Solid lines are functions for response  $R_A$ ; dotted lines are functions for response  $R_B$ . The fixed parameters used to generate the functions were  $n_S = n_T = 5$ ,  $K_A = K_B = 6$ . The channel amplitudes for sustained and transient channels were  $a = 12$  and  $b = 10$ . The channel rates for the sustained channel were (left to right)  $\beta_S(1) = 50$ ,  $\beta_S(2) = 20$ ,  $\beta_S(3) = 10$ , and  $\beta_S(4) = 5$ . The rates for the transient channel were (left to right)  $\beta_T(1) = 10$ ,  $\beta_T(2) = 8$ ,  $\beta_T(3) = 5$ , and  $\beta_T(4) = 3$ .

channels; slower RTs, which occur on trials on which no transient response occurs, are mediated by sustained channels.

## 5. Conclusion

This paper has investigated the properties of a time-inhomogeneous generalization of the Poisson, parallel-counter model of RT, first considered by Pike (1966, 1973) and subsequently developed by Townsend & Ashby (1983). In this generalization the event rates of the Poisson processes that drive the counters are not constant, but are arbitrary functions of time. General expressions were derived for the RT distributions for the time-inhomogeneous model with no restrictions on the Poisson event rates other than that they were continuous. Closed-form expressions for the response probabilities were obtained under a proportional-rates assumption, in which the Poisson event rates were proportional to a single function of time. Closed-form expressions for the mean RTs were obtained for a proportional-rates model in which the integrated event rates were an arbitrary, positive power of time. The value of the time-inhomogeneous version of the model is that it allows various forms of temporal dependencies, of a kind that arise naturally in applications, to be incorporated into the model and their quantitative properties investigated in detail. This in turn makes it possible to formulate quantitative models of a greater richness and generality than has hitherto been possible with models of this class and to assess their goodness of fit using standard statistical techniques.

## Acknowledgements

This research was supported by Australian Research Council Grant A79802778 to Philip Smith and NSF grant SBR-9702291 to Trisha Van Zandt. We thank Douglas Vickers for the hospitality of his laboratory where a part of this work was carried out. We also thank Jim Townsend for helpful discussions on stochastic dominance.

## References

- Audley, R. J. (1960). A stochastic model for individual choice behavior. *Psychological Review*, *67*, 1–15.
- Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice. *British Journal of Mathematical and Statistical Psychology*, *18*, 207–225.
- Burbeck, S. L., & Luce, R. D. (1982). Evidence from auditory simple reaction time for both change and level detectors. *Perception & Psychophysics*, *32*, 117–133.
- Busey, T. A. (1999). Localization and identification rely on different temporal frequencies. *Vision Research*, *39*, 513–532.
- Cox, D. R., & Miller H. D. (1965). *The theory of stochastic processes*. London: Chapman & Hall.
- Cox, D. R., & Oakes, D. (1984). *Analysis of survival data*. London: Chapman & Hall.
- de Lange, H. (1954). Relationship between critical flicker frequency and a set of low frequency characteristics of the eye. *Journal of the Optical Society of America*, *44*, 380–389.
- de Lange, H. (1958). Research into the dynamic nature of the human foveal-cortex systems with intermittent and modulated light: I. Attenuation characteristics with white and colored light. *Journal of the Optical Society of America*, *48*, 777–784.
- Estes, W. K. (1950). Towards a statistical theory of learning. *Psychological Review*, *57*, 94–107.
- Fredericksen, R. E., & Hess, R. F. (1997). Temporal detection in human vision: Dependence on stimulus energy. *Journal of the Optical Society of America, A*, *14*, 2557–2569.

- Gorea, A., & Tyler, C. W. (1986). New look at Bloch's law for contrast. *Journal of the Optical Society of America, A*, 3, 52–61.
- Gradshteyn, I. S., & Ryzhik, I. M. (1965). *Tables of integrals, series, and products*. San Diego, CA: Academic Press.
- Heath, R. A. (1992). A general nonstationary diffusion model for two choice decision making. *Mathematical Social Sciences*, 23, 283–309.
- Kulikowski, J. J., & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient mechanisms in human vision. *Journal of Physiology*, 232, 149–163.
- LaBerge, D. A. (1962). A recruitment model of simple behavior. *Psychometrika*, 27, 375–395.
- LaBerge, D. A. (1994). Quantitative models of attention and response processes in shape identification tasks. *Journal of Mathematical Psychology*, 38, 198–243.
- Legge, G. (1978). Sustained and transient mechanisms in human vision: Temporal and spatial properties. *Vision Research*, 18, 69–82.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–105.
- Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, 103, 603–649.
- Luce R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- 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.
- Merigan, W. H., Byrne, C. E., & Maunsell, J. H. R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience*, 11, 3422–3429.
- Pike, A. R. (1966). Stochastic models of choice behaviour: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, 19, 15–32.
- Pike, A. R. (1973). Response latency models for signal detection. *Psychological Review*, 80, 53–68.
- Purpura, K., Kaplan, E., & Shapley, R. M. (1988). Background light and the contrast gain of primate P and M retinal ganglion cells. *Proceedings of the National Academy of Sciences, USA*, 85, 4534–4537.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Reeves, A. & Sperling, G. (1986). Attention gating in short term visual memory. *Psychological Review*, 93, 180–206.
- Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141–1142.
- Ross, S. M. (1983). *Stochastic processes*. New York: John Wiley.
- Smith, P. L. (1995). Psychophysically-principled models of visual simple reaction time. *Psychological Review*, 102, 567–593.
- Smith, P. L. (1998). Bloch's law predictions from diffusion process models of detection. *Australian Journal of Psychology*, 50, 139–147.
- Smith, P. L. (1998, August). *A stochastic, dynamic theory of visual attention*. Paper presented at the 31st Annual Meeting of the Society for Mathematical Psychology, Vanderbilt University, Nashville, TN.
- Smith, P. L. (in press). Stochastic, dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*.
- Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32, 135–168.
- Smith, P. L. & Vickers, D. (1989). Modeling evidence accumulation with partial loss in expanded judgment. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 797–815.
- Sperling, G. & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America*, 58, 1133–1145.
- Sperling, G. & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review*, 102, 503–532.

- Tolhurst, D. J. (1975a). Sustained and transient channels in human vision. *Vision Research*, *15*, 1151–1155.
- Tolhurst, D. J. (1975b). Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Research*, *15*, 1143–1149.
- Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, *108*, 551–567.
- Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- Valberg, A. & Lee, B. B. (1989). Detection and discrimination of colour, a comparison of physiological and psychophysiological data. *Physica Scripta*, *39*, 178–186.
- Van Zandt, T. (1992). *Modeling nonstationarity in perceptual matching*. Unpublished PhD thesis, Purdue University.
- Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response models applied to perceptual matching. *Psychonomic Bulletin and Review*, *27*, 208–256.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*, 37–58.
- Vickers, D. (1979). *Decision processes in visual perception*. London: Academic Press.
- Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 6.1–6.43). New York: Wiley.
- Watson, A. B. & Nachmias, J. (1977). Patterns of temporal interaction in the detection of gratings. *Vision Research*, *17*, 893–902.

*Received 28 May 1999; revised version received 11 February 2000*