Journal of Mathematical Psychology 97 (2020) 102360

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Can the wrong horse win: The ability of race models to predict fast or slow errors

James T. Townsend, Yanjun Liu *

Indiana University, United States of America

ARTICLE INFO

Article history:
Received 5 December 2019
Received in revised form 11 March 2020
Accepted 31 March 2020
Available online 24 April 2020

Keywords:
Parallel race models
Fast 'right' race models vs. fast 'wrong' race models
Conditional hazard functions for race models
Inhomogeneous Poisson race models
Poisson models with proportional rates

ABSTRACT

This report continues our probe of the fundamental properties of elementary psychological processes. In the present instance, we first distinguish between *descriptive* and *state–space based* parallel race models. Then we show, engaging previous results on stochastic dominance in Theorem 1, that descriptive race models can be designed that predict either faster 'right' channels or faster 'wrong' channels. Moving to state–space based models and in particular, to inhomogeneous Poisson counter models, we use Theorem 1 to prove Theorem 2 which offers sufficient conditions for such models to elicit faster 'rights' than 'wrongs'. Then, constraining ourselves to models possessing proportional processing rates, we revisit an important finding by Smith and Van Zandt (2000) to the effect that in such models, mean processing times conditional on 'right' decisions are faster than those conditional on 'wrong' decisions. Theorem 3 expands that property to the much stronger level of ordered conditional distribution functions. The penultimate section constructs an example of an inhomogeneous Poisson race model that predicts faster 'wrongs' for fast processing times but faster 'rights' for slower processing times. We leave as an open problem the question of whether there exist inhomogeneous Poisson race models where 'wrongs' are stochastically faster than 'rights' for all durations of processing.
© 2020 Elsevier Inc. All rights reserved.

Most mathematical modeling takes place as follows: Theorists and their colleagues construct a model based on specific parameterized stochastic processes or probability distributions. Then they or subsequent researchers conduct experiments which, according to the tenets of the model, produce data which test the model's predictions. This strategy often requires model fits and resulting goodness-of-fit statistics. A rather different style of theorizing proceeds by the theorist investigating properties of entire sets of models (e.g., Marley, 1971), mathematically qualitative methodologies that uncover essential, but non-parametric, characteristics of cognitive processes (Bamber, 1969; Dunn & Kalish, 2018; Townsend, 1990) or classes of models obeying some set of precepts, seeking broad properties, to be evidenced in data, that may distinguish two or more opposing such large sets of models (e.g., Houpt, Townsend, & Jefferson, 2017; Townsend, Wenger, & Houpt, 2018).

The first strategy has the potential advantage that if a model is 'true' (or the closest real-life approximation to truth), and it wins the fit contest, then in principle we are done, accompanied by a new 'law' in science. Of course, the formidable repository of major and widespread instances of model mimicking (e.g., Jones & Dzhafarov, 2014; Khodadadi & Townsend, 2015; Townsend, 1972) already cautions against taking this task lightly. But even beyond that significant general challenge, a specific pitfall is the danger of picking a winner based on specific distributions instead of deeper psychological principles in which scientists are most interested. For example, different hypothetical assumptions concerning the base time can affect the diagnosticity of model selection tools (Townsend & Honey, 2007) and parameter-fitting of a massive amount of choice models (Ratcliff & Tuerlinckx, 2002). Other examples include the observation by Heathcote and Love (2012) that rates and decision thresholds cannot be distinguished when the distribution is lognormal. Thus, a valuable by product of the last mentioned approach is that it encourages the theorist to investigate foundational characteristics of the model classes, which can lead to a deeper comprehension of the relevant dynamics of the competing model types.

Of course, these courses of action are by no means mutually exclusive and they may be and often are, intermingled at various stages of the theoretical enterprise. In fact, we have prescribed an early emphasis on the former to disconfirm large classes of models which embody one position while tentatively affirming its opposition and subsequently move to narrower parameterized models belonging to the affirmed class. This has been referred to as the "sieve approach" to modeling. In any event, it can be instructive to keep the distinction in mind. The present endeavor falls into the second camp. A great deal of our theoretical endeavors has focused on simple architectures and particularly

* Correspondence to: Department of Psychological and Brain Sciences, Indiana University, United States of America.
E-mail address: yanjliu@iu.edu (Y. Liu).

https://doi.org/10.1016/j.jmp.2020.102360
0022-2496/© 2020 Elsevier Inc. All rights reserved.

parallel vs. serial processes and how to test them against one another through, when feasible, distribution and parameter-free predictions (e.g., Algom, Eidels, Hawkins, Jefferson, & Townsend, 2015; Townsend et al., 2018). In other instances, we simply study fundamental properties of these and similar issues (e.g., Balakrishnan, 1994; Houpt et al., 2017; Townsend & Diederich, 1996; Zhang, Liu, & Townsend, 2019). Another issue in response time theory, methodology and practice is error-free (or essentially error-free) vs. measures that include accuracy as well as response times. This investigation is found in the latter domain.

Most stochastic models that can predict errors vs. corrects, or rights vs. wrongs as we will name them in this project, assume that information processing proceeds until a criterion is reached that then produces a decision and response of some kind. Despite our use of *right* and *wrong*, our findings readily apply to arbitrary preference situations (e.g., Marley & Colonius, 1992). These *criterion* models are often called *accumulator models*. That term originally referred to a specific model studied by Smith and Vickers (1988). Since most examples, for instance, diffusion processes, are associated with specific distributional assumptions, we earlier introduced the name *accrual halting models* to include all such models (Townsend, Houpt, & Silbert, 2012). Another reason for the new term is to emphasize that reaching a criterion or decision threshold stops processing. There exist models which can accumulate information, but the stopping process is not governed by hitting a criterion. A dynamic extension of General Recognition Theory by F.G. Ashby is a good example of this class (Ashby, 2000). In his model, there is a processing trajectory in a pattern space which ends at some point in the space. The decision is made according to the 'landing spot' and the response time is determined by the distance of that spot from the nearest decision bound, the so-called "response time distance hypothesis". But all these models assume some sort of state–space that indexes the accrual of information. However, there is a literature, associated especially with multi-sensory-modality perception which utilizes only the notion of the completion time distributions, without reference to underlying state spaces (e.g., Colonius, 1990; Miller, 1978). We call this type of model a "descriptive model". This distinction will be important in our developments.

Townsend and Wenger (1996) have proposed a class of models termed *evidence monitoring models* which assume that while accumulating evidence, the observer possesses a monitor capable of assessing the rate (or amount) of evidence accrual. If the evidence accrual is sufficiently fast or efficient, such processing continues until a decision bound is reached. But, if the accrual is not proceeding sufficiently fast or efficiently, accumulation can cease and a decision can be made on the basis of the accrued information. An analogous model has been proposed by Hawkins and colleagues (Hawkins, Mittner, Forstmann, & Heathcote, 2019) which poses a race between accrual and a guessing process, which would in effect be a special and important case of evidence monitoring theory.

One of the most fundamental characteristics of stochastic models of choice is the issue of whether right responses are faster or slower, or equal, to wrong responses. Empirical evidences have been sustainably found that 'wrong' responses could be sometimes faster or slower than 'right' responses, for instance in visual discrimination tasks where discriminations are easier (e.g., Swensson, 1972) and in certain motion discrimination paradigms (e.g., Ratcliff & McKoon, 2008). The famous *sequential probability ratio test* model, though appealing on many dimensions, is notorious for predicting equal RTs for the two decisions (Laming, 1973; Ratcliff, 1978; Stone, 1960; Townsend & Ashby, 1983). Various strategies can be engaged to surmount this conundrum. For instance, by implementing the moment generating function for increments to the random walk, response conditioned mean times of 'right' and 'wrong' were no longer forced to be equal (Link & Heath, 1975). Ditterich developed a time-variant version of the diffusion model allowing a gain of the sensory signals to increase over time which can account for longer error response times (Ditterich, 2006a, 2006b). The best known, due to Ratcliff and colleagues (Ratcliff, 1978; Smith & Ratcliff, 2004), has been to install distributions on the drift rate and the starting point or, equivalently, the decision criteria (e.g., Wagenmakers, Ratcliff, Gomez, & McKoon, 2008). The resulting probability mixtures can deliver either faster 'rights' or faster 'wrongs'.

Current accrual halting models which assume right and wrong information is accrued simultaneously, tend to take the form of either a parallel race among the n possibilities (e.g., Smith & Van Zandt, 2000; Townsend & Ashby, 1983), or a random walk or diffusion process (e.g., Busemeyer & Townsend, 1993; Link & Heath, 1975; Ratcliff, 1978), though the preponderance has been with the latter tack. On the other hand,descriptive race models, as mentioned above, are formulated simply as a distribution on the minimum processing time (e.g., Audley, 1973; Bundesen, 1990; Colonius, 1990; Colonius & Vorberg, 1994; Marley & Colonius, 1992; Raab, 1962; Townsend & Ashby, 1983).

The first, descriptive case is depicted in Fig. 1a which shows the wrong process winning thus producing an error. The second accrual halting type, seen in Fig. 1b, exhibits a more detailed account based on a race between two Poisson counters with the same decision threshold. Both can contribute to our knowledge of psychological systems and both will, indeed, play a role in the present investigation. We also observe before proceeding that random walk and diffusion models can be viewed as a limiting case of parallel channels which interact in a mutually inhibitory fashion (e.g., Colonius & Townsend, 1997).

The present focus will be on parallel, minimum-time models, in the usual sense where the channels are in a race to the finish, and it will be further assumed that: 1. The parallel channels are stochastically independent. 2. One channel is operating on the 'wrong' input. One channel is operating on the 'right' input. 3. The winner of a race determines the choice and response. 4. It will always be assumed that the 'right' stimulus is processed at a faster rate (to be made precise subsequently) than the 'wrong' stimulus. 5. Both channels have the same response criterion so there is no response bias assumed here. 6. There are very few actual applications of these models to $n > 2$ and we shall concentrate on $n = 2$, but the simplicity of independent race models suggests ready generalization of our results.

Now, let $T =$ processing time and let t be a specific value of T. Notations of symbols that are utilized in the rest of paper are summarized in Table 1. With respect to independent races, our present targeted models, it might be intuitively conjectured that because 'rights' have to be faster than 'wrongs' in order that $P(R) > \frac{1}{2}$, processing time conditioned on 'right' responses will also be faster in the usual experimental sense: $P(T \leq t|R)$. However, our theoretical inquiry uncovers the rather astonishing result that stochastic race models need not, in fact, force this result.

Moreover, we also probe another essential characteristic of information processing: Whether in independent race models, the probability of being 'right' conditional on the processing time t, increases as a function of t or not. That is, is $P(R|t)$ an increasing function of t? As it happens, this property, like that of whether 'wrongs' are always slower in a stochastic sense, than 'rights', is not always true. That is, there exist independent race models where $P(R|t)$ is a decreasing function of t.

We next have to decide at what level to formulate our investigation. For instance, if 'rights' are faster than 'wrongs', at what level does that occur (see e.g., Townsend, 1990; Townsend &

Table 1
Summary of notations.

Symbols	Meaning
T	A random variable of total processing time
t	A numerical value of T that is nonnegative
i	a specific response taking values of either $R = right$ or $W = wrong$.
T_i	A random channel processing time
$P(T \leq t\|i)$	The cumulative distribution function of processing time conditional on response i
$P(i\|t)$	Probability of being 'right' or 'wrong' conditional on a specific value of total processing time
$E(T\|i)$	Mean of the processing time conditional on 'right' or 'wrong'
$g_i(t)$	Probability density function of channel processing time for 'right' or 'wrong'
$h_i(t)$	Hazard function of channel processing time
$r(t)$	Ratio of hazard functions of channel processing time
$h(t\|i)$	Hazard function of total processing time conditional on response

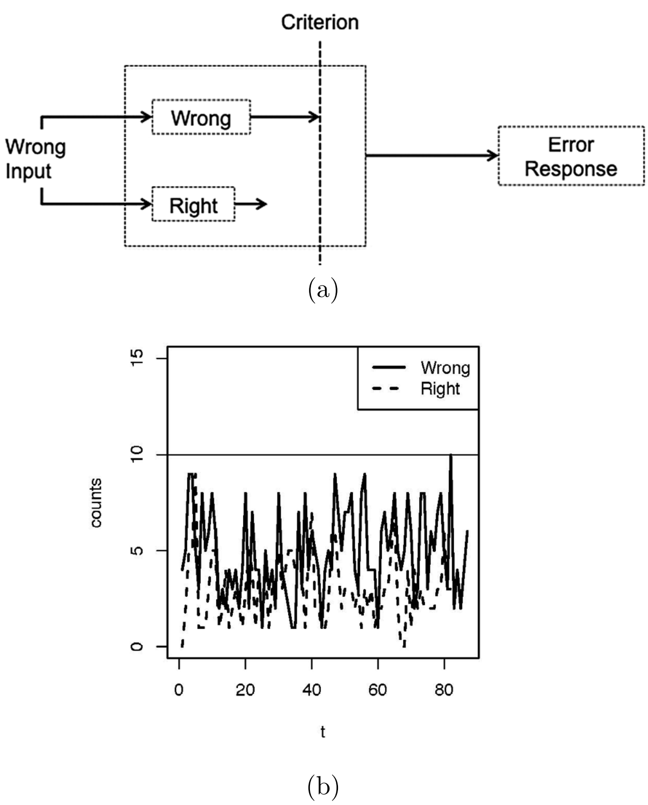

Fig. 1. (a) An example of race model where the 'wrong' signal has been input and 'wrong' wins. (b) An example model of two Poisson counters where 'wrong' hits the criterion = 10 first and produces an error response. t has arbitrary unit.

Ashby, 1978)? Perhaps the ordering only holds for mean RTs, that is for, say $E(T|R)$ vs. $E(T|W)$ (where E, the expectation operator, is equivalent to the arithmetic mean). Knowledge at this level is quite useful of course, but it would be of import to explore the possibility of statistical orderings at stronger levels. For instance, Townsend (1990) argues that finding an ordering say, such that $P(T \leq t|R) < P(T \leq t|W)$ (or vice versa) makes a much stronger statement about nature than a finding only at the level of means, and in fact, it happens that the former order implies a ordering of the means as well as the medians. In point of fact, we will demonstrate that descriptive race models exist that can predict either faster processing times when conditioning on 'wrongs' than when conditioning on 'rights' or the opposite. Similarly, these results are attained at a strong level in the Townsend (1990) scheme, even while maintaining the postulate that rights have to win the race more than half of the time. Theorem 1 will prove this assertion and, in the process, illustrate how hazard functions govern the relative speeds of the two channels.

After dealing only with the minimum time statistic of two independent distributions, we expand our purview to analyze accrual halting channels interpretation of the race. An especially important set of such models, still quite general, is that founded on inhomogeneous Poisson counters, which emphasizes the counting process or, equivalently, inhomogeneous Gamma processes for the completion times. This type of model was thoroughly investigated by Smith and Van Zandt (2000). Here, each intercompletion time (Townsend & Ashby, 1983, duration between two successive completions; see, e.g.,) is distributed by a generalized exponential where the rate is itself a function of time. Moreover, there is a special subclass of this kind of model wherein the processing (i.e., counting) rates are proportional. In our context, this would mean that the 'right' gamma rates are proportional to the 'wrong' rates. Smith and Van Zandt (2000) made the important discovery that proportional-rates, inhomogeneous, right–wrong race models predict that the mean processing times, conditional on being 'right' are inevitably faster than those conditional on being 'wrong'. With that prediction in mind, our Theorem 2 moves ahead to construct sufficient conditions for 'right' processing times to be faster than 'wrong' processing times in the stronger sense that the probability that a 'right' finishing time is faster than some arbitrary time t, will always be greater than the likelihood that a 'wrong' finishing time will be faster than that time t.

More critically, when we specialize to the class explored by Smith and Van Zandt (2000), we discover from Theorem 3 that proportional-rates, inhomogeneous Gamma race models (that is, the same category for which Smith and Van Zandt proved their results for the conditional mean processing times) fulfill the sufficient conditions of Theorem 2. Therefore, they always elicit faster 'rights' than 'wrongs' at the strong distributional level of that theorem. Thus, we determine that, indeed, at least race models of a certain general breed, are inclined to predict faster 'rights' than 'wrongs' and that at a quite robust stochastic echelon, thereby fortifying and expanding the Smith and Van Zandt (2000) conclusions.

In short, the approach we take in this paper is first exploring the principle of conditioned processing times of 'right' and 'wrong' with a broad class of race models that are free of any distributional assumption (see Theorem 1), that is descriptive race models. Then, we examine these general findings with a subclass of race models that involve time-varying Poisson counting processes (see Theorem 2), and finalize our theoretical exploration with a special case of this Poisson counting model assuming proportional counting rates (see Theorem 3).

The characteristics of our broad class of race models, as given informally above, can be taken as axioms and are not put down formally since they are so well-known and accepted. However, we should state an important further assumption. As given in Townsend (1990), 'rights' decisions could be faster or slower than 'wrong' decisions at a number of hierarchical stations. So, we first require an axiom on hazard functions to ensure that more 'rights' are made than 'wrongs'.

1. Axiom on hazard functions of right vs. wrong channels

Let the hazard function for the 'wrong' channel be $h_W(t) = g_W(t)/\overline{G_W}(t)$, where $g_W(t)$ is the probability density function on 'wrong' processing times, $\overline{G_W}(t) = 1 - G_W(t)$ and $G_W(t)$ is the cumulative distribution function on 'wrong' processing times. $\overline{G_W}(t)$ is just the so-called *survivor function* of the channel processing time. Similarly, the hazard function for the 'right' channel is $h_R(t) = g_R(t)/\overline{G_R}(t)$ with the analogous interpretation. Then we assume that $h_R(t) > h_W(t)$, for all $t \geq 0$.

The consequent lemma insures that the 'rights' frequency exceeds the number of 'wrongs'.

Lemma 1. *Because $h_R(t) > h_W(t)$, it follows that $P(R) > P(W)$.*

Proof. See Townsend (1990). □

Again, from the discussion above, one might intuit that because 'right' channels operate more speedily than 'wrong' channels, that conditioning on a 'right' decision would always find these faster than 'wrong' decisions. This intuition would be misguided as we shall shortly bear witness.

In fact, it intriguingly turns out that the behavior of the ratios of the hazard functions over time determines whether 'rights' win the race faster than 'wrongs' as we see in the following theorem. From here on out, we will use the expression *strictly increasing* (or *decreasing*) to mean *strictly monotonically increasing* (or *decreasing*). We will also need the new conditional, winning vs. losing, densities $f(t|R) = P(T_R = t \leq T_W|R) = \frac{g_R(t)\overline{G_W}(t)}{P(R)}$ and $f(t|W) = P(T_W = t \leq T_R|W) = \frac{g_W(t)\overline{G_R}(t)}{P(W)}$. Similarly, we can readily determine the respective 'right' and 'wrong' conditional hazard functions to be $h(t|R) = \frac{f(t|R)}{\overline{F}(t|R)}$ and $h(t|W) = \frac{f(t|W)}{\overline{F}(t|W)}$, where $F(t|R)$ and $F(t|W)$ are cumulative distributions of total processing time conditional on 'right' and 'wrong', respectively. The reader will notice that in a felicitous manner, the terms $P(R)$ and $P(W)$ cancel in the hazard functions so that $h(t|R) = \frac{g_R(t)\overline{G_W}(t)}{\int_t^\infty g_R(s)\overline{G_W}(s)ds}$ and $h(t|W) = \frac{g_W(t)\overline{G_R}(t)}{\int_t^\infty g_W(s)\overline{G_R}(s)ds}$.

Theorem 1. *Consider the ratio of the terms 'right' wins the race at time t vs. 'wrong' wins the race at time t: $r(t) = \frac{g_R(t)/\overline{G_R}(t)}{g_W(t)/\overline{G_W}(t)} = \frac{g_R(t)\overline{G_W}(t)}{g_W(t)\overline{G_R}(t)}$. If $r(t)$ strictly decreases, then errors are stochastically slower than corrects in the sense that $P(T \leq t|R) > P(T \leq t|W)$. But if it strictly increases, then the reverse ordering occurs. Observe that $r(t)$ is also the ratio $\frac{h_R(t)}{h_W(t)}$, that is the ratio of the 'right' to 'wrong' hazard functions.*

Proof. $P(T \leq t|R) = \int_0^t \frac{g_R(s)\overline{G_W}(s)}{P(R)} ds = F(t|R)$ is the conditional cumulative distribution on 'right' ; whereas $P(T \leq t|W) = \int_0^t \frac{g_W(s)\overline{G_R}(s)}{P(W)} ds = F(t|W)$ is the conditional cumulative distribution on 'wrong'. We first consider the case when $r(t)$ is monotonically decreasing. Following the method used in Townsend and Ashby (1983), we have

$$r(t)\int_t^\infty g_W(s)\overline{G_R}(s)ds > \int_t^\infty r(s)g_W(s)\overline{G_R}(s)ds$$

$$r(t)g_W(t)\overline{G_R}(t)\int_t^\infty g_W(s)\overline{G_R}(s)ds > g_W(t)\overline{G_R}(t) \times \int_t^\infty r(s)g_W(s)\overline{G_R}(s)ds$$

$$\frac{r(t)g_W(t)\overline{G_R}(t)}{\int_t^\infty r(s)g_W(s)\overline{G_R}(s)ds} > \frac{g_W(t)\overline{G_R}(t)}{\int_t^\infty g_W(s)\overline{G_R}(s)ds}$$

$$\frac{g_R(t)\overline{G_W}(t)}{\int_t^\infty g_R(s)\overline{G_W}(s)ds} > \frac{g_W(t)\overline{G_R}(t)}{\int_t^\infty g_W(s)\overline{G_R}(s)ds}$$

$$h(t|R) > h(t|W)$$

Since the conditional hazard functions of 'right' and 'wrong' are ordered, by Proposition 8 in Townsend (1990), it follows that $F(t|R) > F(t|W)$, which is equivalent to $P(T \leq t|R) > P(T \leq t|W)$.

Following similar logic, it is clear to see that when $r(t)$ is monotonically increasing, then $P(T \leq t|R) < P(T \leq t|W)$. □

Corollary 1.1. *In the event that $r(t)$ decreases, then 'rights' are faster than 'wrongs' in the sense that $P(T \leq t|R) > P(T \leq t|W)$ and also the mean and median processing time for 'rights' are smaller than those for 'wrongs'. Conversely, if $r(t)$ increases, 'wrongs' are faster than 'rights' at the distributional level as well as the conditional means and medians level.*

Proof. Proposition 4 in Townsend (1990) states that the ordering on cumulative distributions implies the same ordering at mean and median levels. Since the monotonicity of $r(t)$ implies ordered conditional cumulative distributions here, it thereby ensures this conclusion. □

The behavior of 'rights' being faster than 'wrong' may seem somewhat pedestrian given the earlier, even if misguided, intuition. It is more intriguing to discover rather elementary distributions that actually produce faster 'wrongs' than 'rights'. In fact, we can unearth quite simple examples of either behavior within the class of hazard functions defined as affine functions of time. That is, $h(t) = at + b$. Note that this simple function can be viewed a very special case of inhomogeneous (time varying) gamma processes, since the time-varying rate is just the hazard function $h(t)$, for an intercompletion time in such a stochastic process. The family density function is then simply $g(t) = (at + b)e^{(\frac{-at^2}{2} + bt)}$.

We begin with the mundane set of circumstances where $r(t)$ decreases, which elicits faster 'rights'. Suppose $h_R(t) = 4t + 2.1$ while $h_W(t) = 2t + 1$. Notice that $h_R(t) > h_W(t)$ as expected and then calculate the first derivative of the ratio function: $\frac{dr(t)}{dt} = \frac{-0.2}{(2t+1)^2}$ which is always negative for $t \geq 0$. Thus, $r(t)$ is strictly decreasing. Fig. 2a shows the associated density functions. Fig. 2b shows $r(t)$. Fig. 2c reveals the two conditional hazard functions and Fig. 2d shows the difference of conditional functions with respect to t: $P(T \leq t|R) - P(T \leq t|W)$.

Next, take the new hazard functions $h_R(t) = (7t + 2) > h_W(t) = 3t + 1$. Again, 'rights' are more frequent than 'wrongs' but now $r(t)$ increases thus certifying that 'wrongs' are now faster than 'rights'. Fig. 3a, b, c, d graph the functions corresponding to those in Fig. 2, therefore illustrate the opposite kind of behavior from the case where $r(t)$ decreases.

The last result of this section fulfills the earlier promise to shed some light on $P(R|t)$ which we exhibit in Corollary 1.2. The reader may be mildly shocked, as we were, to see how similar all the distributional characteristics look, except for the likelihood ratio and the difference in conditional distribution functions.

Corollary 1.2. *If $r(t)$ decreases, then $P(R|T = t)$ decreases as a function of t and if $r(t)$ increases, $P(R|T = t)$ increases as a function of t.*

Proof.

$$P(R|T = t) = \frac{P(R, T = t)}{P(T = t)} = \frac{g_R(t)\overline{G_W}(t)}{g_R(t)\overline{G_W}(t) + g_W(t)\overline{G_R}(t)} = \frac{1}{1 + \frac{1}{r(t)}}$$

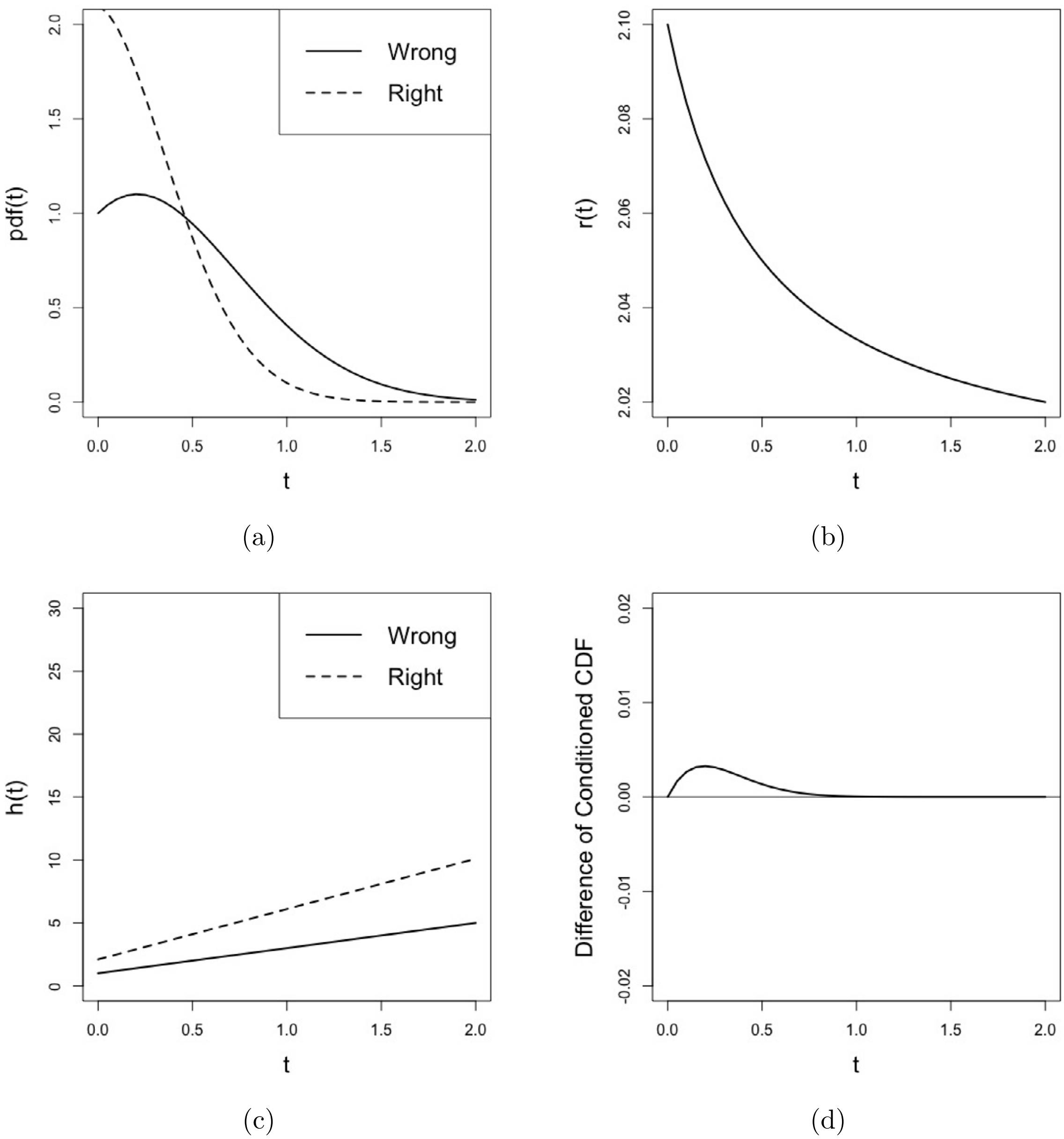

Fig. 2. (a) Density functions of RTs for 'right' and 'wrong' channels. (b) Decreasing ratio of hazard functions, $r(t) = \frac{h_R(t)}{h_W(t)}$. (c) Correspondent hazard functions of 'right' and 'wrong' channels. (d) Difference of conditioned distributions: $P(T \leq t|R) - P(T \leq t|W)$. Notice that differences are above zero for $t > 0$, indicating that 'wrong' responses are stochastically slower than 'right' responses.

The finding ensues that $P(R|t)$ moves in the direction of $r(t)$ with respect of t. □

Thus, we have an intriguing conjoining of predictions: If the likelihood of being right grows with t ($r(t)$ is increasing) then wrongs should be faster than rights and if being right declines with t ($r(t)$ decreases over time) then rights are predicted to be faster than wrongs. It would be of interest to probe this connection with other models, for example, diffusion processes. Smith (personal communication, 2020) derived the equations of conditional accuracy functions on current state for Wiener process. Results show that when there is cross-variability in drift rates, the conditional probability of accuracy decreases with t; whereas when drift rates are not cross-trial varied, probabilities of conditional accuracy are independent of t. His results would correspond to what might occur if we conditioned on number of counts in the two channels at an arbitrary time. This path could be quite interesting but lies beyond our present scope.

2. Race models based on time inhomogeneous Poisson rates, proportional rates models, and standard gamma races: Faster rights, slower wrongs

Poisson counting models with their twin processing time distributions, the gamma waiting time densities, have historically been highly popular among theorists (e.g., Luce, 1986; McGill, 1963; McGill & Gibbon, 1965; Smith & Van Zandt, 2000; Townsend & Ashby, 1983). The general gamma distributions, elicited by letting the processing rates differ among channels or items have also proven valuable though not so prevalent as the ordinary variety (e.g., Colonius & Vorberg, 1994; McGill, 1963; Townsend & Ashby, 1983).

As noted earlier, a different strategy of generalization of descriptive and predictive power can be brought to bear by letting processing rates vary across time, thus evoking the class of time-dependent (or inhomogeneous) Poisson processes, again with their twins, the time-varying (also called inhomogeneous)

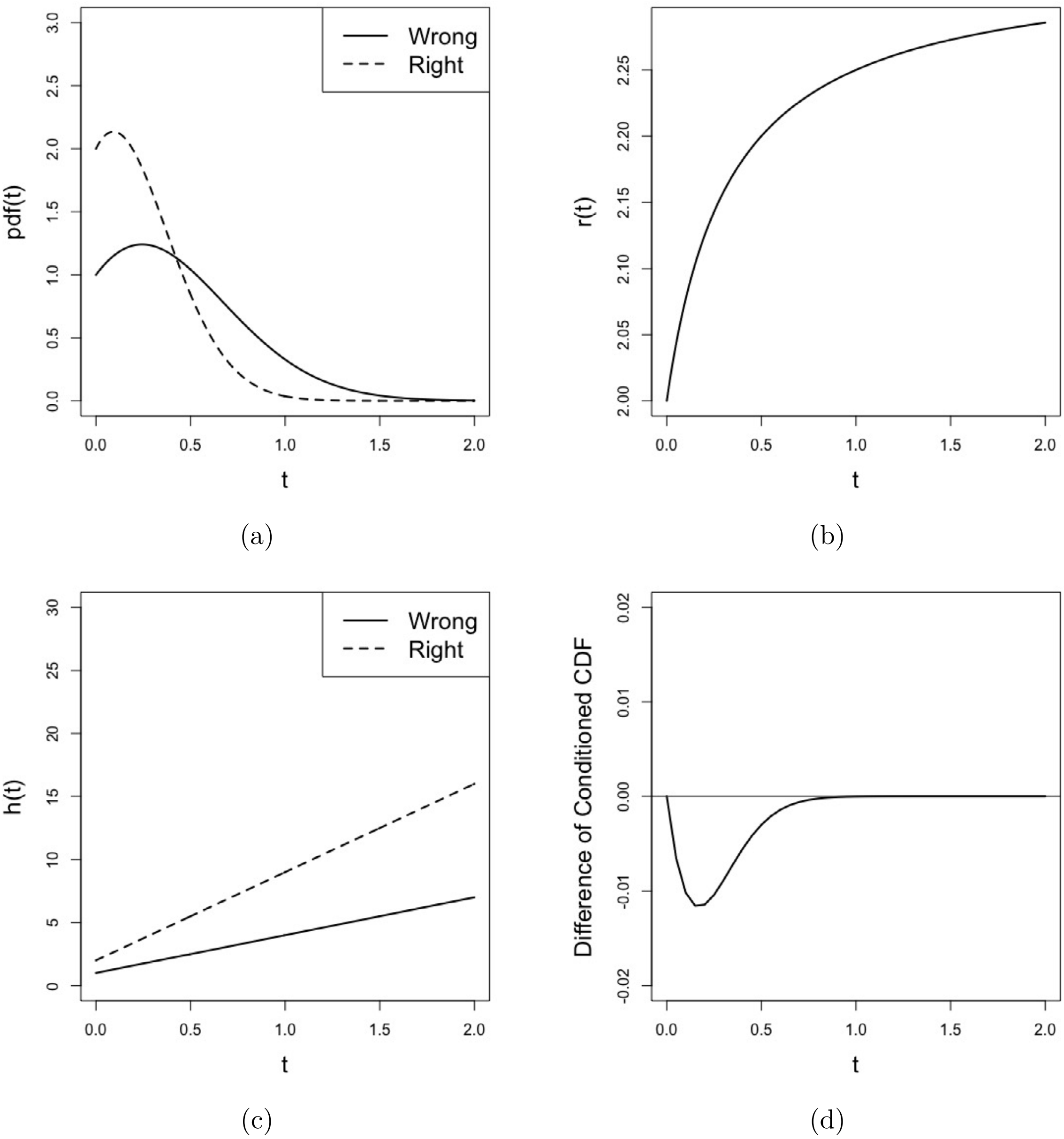

Fig. 3. (a) Density functions of RTs for 'right' and 'wrong' channels. (b) Increasing ratio of hazard functions, $r(t) = \frac{h_R(t)}{h_W(t)}$. (c) Correspondent hazard functions of 'right' and 'wrong' channels. (d) Difference of conditioned distributions: $P(T \leq t|R) - P(T \leq t|W)$. Notice that differences are below zero for $t > 0$, indicating that 'wrong' responses are stochastically faster than 'right' responses.

Gamma distributions. Of course, the classical special case for this class resides in the standard gamma process. Smith and Van Zandt (2000) contributed a landmark study on time-dependent Poisson processes. One of their many findings was on the important special case where the 'right' channel's rate parameter is proportional to the 'wrong' channel's rate parameter. That is, $R(t) = \alpha W(t)$, with α being a number greater than 1. Then, they proved that 'rights' are faster than 'wrongs' at the conditional mean processing time level. Our next demonstration enlists sufficient conditions to deliver the stronger stochastic ordering on conditional hazard functions for general time-varying Poisson counter processes. As above, this exercise leads to the strong stochastic relationship: $P(T \leq t|R) > P(T \leq t|W)$, that is, it is always more likely that 'rights' finish faster than 'wrongs'.

Although in some sense, the next finding is a corollary of Theorem 1, because we deem it sufficiently valuable, we grace it with the designation 'theorem'. The targeted models are those within the class of general inhomogeneous Gamma processes. Let $R'(t)$ be the time-derivative of $R(t)$ and $W'(t)$ the time derivative of $W(t)$.

Theorem 2. *Let $W(t)$ be the time-variable rate parameter for the 'wrong' channel and $R(t)$ be the time-variable rate parameter for the 'right' channel. Let k stages be required for a decision on either channel . Next assume following Axioms hold: 1. $R'(t) > W'(t) > 0$. 2. $\frac{R(t)}{W(t)}$ is decreasing in t. 3. $\frac{R'(t)}{W'(t)}$ is decreasing in t. 4. $\frac{R'(t)}{W'(t)} > \frac{\sum_{j=0}^{k-1} \frac{R^j(t)}{j!} \sum_{i=0}^{k-2} \frac{W^i(t)}{i!}}{\sum_{j=0}^{k-1} \frac{W^j(t)}{j!} \sum_{i=0}^{k-2} \frac{R^i(t)}{i!}}$. Then by Theorem 1, 'rights' are stochastically faster than 'wrongs'.*

Proof. Define the probability density function of the time-varying Poisson process for the 'wrong' and 'right' channels as following

Smith and Van Zandt (2000):

$$g_W(t) = \frac{W(t)^{k-1}W'(t)e^{-W(t)}}{(k-1)!}$$
$$g_R(t) = \frac{R(t)^{k-1}R'(t)e^{-R(t)}}{(k-1)!} \tag{1}$$

Define the survivor function of the time-varying Poisson process for the 'wrong' and 'right' channels as following Smith and Van Zandt (2000):

$$\overline{G_W}(t) = \sum_{j=0}^{k-1} \frac{W(t)^j}{j!} e^{-W(t)}$$
$$\overline{G_R}(t) = \sum_{j=0}^{k-1} \frac{R(t)^j}{j!} e^{-R(t)} \tag{2}$$

Then,

$$\begin{aligned} r(t) &= \frac{g_R(t)\overline{G_W}(t)}{g_W(t)\overline{G_R}(t)} \\ &= \frac{R(t)^{k-1}R'(t)e^{-R(t)}\sum_{j=0}^{k-1}\frac{W(t)^j}{j!}e^{-W(t)}}{W(t)^{k-1}W'(t)e^{-W(t)}\sum_{j=0}^{k-1}\frac{R(t)^j}{j!}e^{-R(t)}} \\ &= \frac{R(t)^{k-1}R'(t)\sum_{j=0}^{k-1}\frac{W(t)^j}{j!}}{W(t)^{k-1}W'(t)\sum_{j=0}^{k-1}\frac{R(t)^j}{j!}} \end{aligned} \tag{3}$$

We employ the differentiation product rule to show that Axioms 1, 2, 3 are sufficient to prove the claim. Given Axioms 1, 2, 3, $\frac{R(t)^{k-1}R'(t)}{W(t)^{k-1}W'(t)}$ is decreasing as t increases. Given Axiom 4, we have

$$\begin{aligned} &\frac{d\left[\frac{\sum_{j=0}^{k-1}\frac{W^j(t)}{j!}}{\sum_{j=0}^{k-1}\frac{R^j(t)}{j!}}\right]}{dt} \\ &= \sum_{j=0}^{k-1}\frac{R^j(t)}{j!}W'(t)\sum_{i=0}^{k-2}\frac{W^i(t)}{i!} - \sum_{j=0}^{k-1}\frac{W^j(t)}{j!}R'(t)\sum_{i=0}^{k-2}\frac{R^i(t)}{i!} \\ &= W'(t)\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{R^j(t)}{j!}\frac{W^i(t)}{i!} - R'(t)\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{W^j(t)}{j!}\frac{R^i(t)}{i!} \\ &< 0 \end{aligned}$$

Thus, $\frac{\sum_{j=0}^{k-1}\frac{W(t)^j}{j!}}{\sum_{j=0}^{k-1}\frac{R(t)^j}{j!}}$ decreases as t increases.

Combining the above results, Axioms 1, 2, 3 and 4 ensure that $r(t)$ is a decreasing function of t. According to Theorem 1, we can conclude that such time-varying Poisson process produces stochastically faster 'rights' than 'wrongs'. □

Satisfying the inequalities in Theorem 2 therefore generates stochastic dominance in processing times of 'rights' vs. 'wrongs'. The particular case of inhomogeneous Gamma processes where the 'right' rates (as functions of time) are proportional to the 'wrong' rates of especial interest (e.g., Cox, 1972; Lappin & Eriksen, 1966; Smith & Van Zandt, 2000; Wenger & Townsend, 2000). Smith and Van Zandt (2000) found that, for these models, the assumption of $R(t) = \alpha W(t), \alpha > 1$ was sufficient to elicit the results that, letting T be the random variable designating processing time, $E[T|R] < E[T|W]$. Our assumptions are more stringent as is visible in the axioms of Theorem 2. However, in the special case of inhomogeneous Gamma processes with proportional rates, we acquire Axioms 1, 2, 3 free so to speak so we only need to prove that essential Axiom 4 is true. But this stipulation is then reduced to that of Smith and Van Zandt (2000).

Hence, because we can then demonstrate the veracity of Axiom 4 in the next theorem, we learn that their assumption forces shorter processing time of 'rights' over 'wrongs' at a considerably stronger stochastic level.

Theorem 3. *Consider the proportional rates inhomogeneous Poisson counting models with Gamma distributions. Then we proceed to show that if Axioms 1, 2, 3 of Theorem 2 hold–then the critical Axiom 4 holds.*

Proof. Recall the operative formula: $r(t) = \frac{g_R(t)\overline{G_W}(t)}{g_W(t)\overline{G_R}(t)}$ which must be demonstrated to decrease when interpreted as an inhomogeneous Gamma process. We see that for a time-varying Poisson counting model with decision criterion of k stages (Eq. (3)): $r(t) = \frac{g_R(t)\overline{G_W}(t)}{g_W(t)\overline{G_R}(t)} = [\frac{R'(t)}{W'(t)}][\frac{R^{k-1}(t)}{W^{k-1}(t)}[\frac{e^{-R(t)}\sum_{j=0}^{k-1}\frac{W^j(t)e^{-W(t)}}{j!}}{e^{-W(t)}\sum_{j=0}^{k-1}\frac{R^j(t)e^{-R(t)}}{j!}}]]$.

After canceling the exponential terms, by the product rule, we now show that, given proportional processing rates $R(t) = \alpha W(t)$, the Axioms 1, 2, 3 will suffice to ensure Axiom 4 holds and thus Theorem 1 is in force.

$$\begin{aligned} &\frac{d\left[\frac{\sum_{j=0}^{k-1}\frac{W^j(t)}{j!}}{\sum_{j=0}^{k-1}\frac{R^j(t)}{j!}}\right]}{dt} \\ &= \sum_{j=0}^{k-1}\frac{R^j(t)}{j!}W'(t)\sum_{i=0}^{k-2}\frac{W^i(t)}{i!} - \sum_{j=0}^{k-1}\frac{W^j(t)}{j!}R'(t)\sum_{i=0}^{k-2}\frac{R^i(t)}{i!} \\ &= W'(t)\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{R^j(t)}{j!}\frac{W^i(t)}{i!} - R'(t)\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{W^j(t)}{j!}\frac{R^i(t)}{i!} \end{aligned}$$

and now introducing the proportional rates, $R(t) = \alpha W(t)$, where we assume as usual $R(t), W(t) > 0$, for $t \geq 0$ (as well of course as Axioms 1, 2, 3), $\alpha > 1$. Then above equation turns to be

$$W'(t)[\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{\alpha^j W^{i+j}(t)}{j!i!} - \alpha\sum_{j=0}^{k-1}\sum_{i=0}^{k-2}\frac{\alpha^i W^{i+j}(t)}{j!i!}] \tag{4}$$

Next it is useful to set $i + j = m$ and separately consider two cases; case I: $m \leq k-2$; case II: $k-1 \leq m \leq 2k-3$.

Case I: let $i + j$ be equal to arbitrary m, $m \leq k-2$.

$$\begin{aligned} (4) &= W'(t)[\sum_{x=0}^{m}\frac{\alpha^x W^m(t)}{(x)!(m-x)!} - \alpha\sum_{y=0}^{m}\frac{\alpha^{m-y}W^m(t)}{(m-y)!(y)!}] \\ &= W'(t)W^m(t)[\sum_{x=0}^{m}\frac{\alpha^x}{(x)!(m-x)!} - \alpha\sum_{y=0}^{m}\frac{\alpha^{m-y}}{(m-y)!(y)!}] \end{aligned} \tag{5}$$

Substitute y with $m - x$, where x = 0, 1, 2, …, m-1, m and y = m, m-1, …, 2, 1, 0.

$$(5) = W'(t)W^m(t)[\sum_{x=0}^{m}\frac{\alpha^x}{(x)!(m-x)!}(1-\alpha)] < 0 \text{ , since } \alpha > 1.$$

Case 2: let $i + j$ be equal to arbitrary $m = k - 1 + a$, $k - 1 \leq m \leq 2k - 3$ and $0 \leq a \leq k - 2$.

$$\begin{aligned} (4) &= W'(t)W^{k-1+a}(t)[\frac{\alpha^{k-1}}{(k-1)!a!} + \frac{\alpha^{k-2}}{(k-2)!(a+1)!} + \cdots \\ &+ \frac{\alpha^{a+2}}{(a+2)!(k-3)!} + \frac{\alpha^{a+1}}{(a+1)!(k-2)!}] \\ &- \alpha W'(t)W^{k-1+a}(t)[\frac{\alpha^a}{(k-1)!a!} + \frac{\alpha^{a+1}}{(k-2)!(a+1)!} + \cdots \\ &+ \frac{\alpha^{k-3}}{(a+2)!(k-3)!} + \frac{\alpha^{k-2}}{(a+1)!(k-2)!}] \end{aligned}$$

$$= W'(t)W^{k-1+a}(t)[\sum_{x=0}^{k-2-a}\frac{\alpha^{k-1-x}}{(k-1-x)!(a+x)!} - \alpha\sum_{y=0}^{k-2-a}\frac{\alpha^{a+y}}{(k-1-y)!(a+y)!}] \quad (6)$$

Substitute y with $k-a-x-2$, where $x = 0, 1, 2, \ldots, k-3-a, k-2-a$ and $y = k-2-a, k-3-a, \ldots, 2, 1, 0$

$$(6) = W'(t)W^{k-1+a}(t)[\sum_{x=0}^{k-2-a}\frac{\alpha^{k-1-x}}{(k-1-x)!(a+x)!} - \sum_{x=0}^{k-2-a}\frac{\alpha^{k-1-x}}{(k-2-x)!(a+x+1)!}]$$
$$= W'(t)W^{k-1+a}(t) \times [\sum_{x=0}^{k-2-a}\frac{(k-2-x)!(a+1+x)!-(k-1-x)!(a+x)!}{(k-1-x)!(a+x)!(k-2-x)!(a+x+1)!} \times \alpha^{k-1-x}] \quad (7)$$

We have total $k-1-a$ number of terms in the above geometric summation.

First, let us consider when $k-1-a$ is even. Then,

$$(7) = W'(t)W^{k-1+a}(t)[\frac{(a+1)!(k-2)!-(a)!(k-1)!}{(a+1)!(k-2)!(a)!(k-1)!}\alpha^{k-1} + \frac{(a+2)!(k-3)!-(a+1)!(k-2)!}{(a+2)!(k-3)!(a+1)!(k-2)!}\alpha^{k-2} + \cdots + \frac{(\frac{a+k-1}{2})!(\frac{k+a-1}{2})!-(\frac{a+k-3}{2})!(\frac{k+a+1}{2})!}{(\frac{a+k-1}{2})!(\frac{k+a-1}{2})!(\frac{a+k-3}{2})!(\frac{k+a+1}{2})!}\alpha^{\frac{k+a+1}{2}} + \frac{(\frac{a+k+1}{2})!(\frac{k+a-3}{2})!-(\frac{a+k-1}{2})!(\frac{k+a-1}{2})!}{(\frac{a+k+1}{2})!(\frac{k+a-3}{2})!(\frac{a+k-1}{2})!(\frac{k+a-1}{2})!}\alpha^{\frac{k+a-1}{2}} + \cdots + \frac{(k-2)!(a+1)!-(k-3)!(a+2)!}{(k-2)!(a+1)!(k-3)!(a+2)!}\alpha^{a+2} + \frac{(k-1)!(a)!-(k-2)!(a+1)!}{(k-1)!(a)!(k-2)!(a+1)!}\alpha^{a+1}].$$

After combining like terms, above equation turns to be:

$$(7) = W'(t)W^{k-1+a}(t) \times [\sum_{x=0}^{\frac{k-a-3}{2}}\frac{(k-2-x)!(a+1+x)!-(k-1-x)!(a+x)!}{(k-1-x)!(a+x)!(k-2-x)!(a+x+1)!} \times (\alpha^{k-1-x}-\alpha^{a+1+x})]$$
$$< 0 \text{ , since } 0 \le a \le k-2 \text{ and } \alpha > 1.$$

Now, let us consider when $k-1-a$ is odd. It follows that

$$(7) = W'(t)W^{k-1+a}(t)[\frac{(a+1)!(k-2)!-(a)!(k-1)!}{(a+1)!(k-2)!(a)!(k-1)!}\alpha^{k-1} + \frac{(a+2)!(k-3)!-(a+1)!(k-2)!}{(a+2)!(k-3)!(a+1)!(k-2)!}\alpha^{k-2} + \cdots + \frac{(\frac{a+k-2}{2})!(\frac{k+a}{2})!-(\frac{a+k-4}{2})!(\frac{k+a+2}{2})!}{(\frac{a+k-2}{2})!(\frac{k+a}{2})!(\frac{a+k-4}{2})!(\frac{k+a+2}{2})!}\alpha^{\frac{k+a+2}{2}} + \frac{(\frac{a+k}{2})!(\frac{k+a-2}{2})!-(\frac{k-2+a}{2})!(\frac{k+a}{2})!}{(\frac{a+k}{2})!(\frac{k+a-2}{2})!(\frac{k-2+a}{2})!(\frac{k+a}{2})!}\alpha^{\frac{k+a}{2}} + \frac{(\frac{a+k+2}{2})!(\frac{k+a-4}{2})!-(\frac{k+a}{2})!(\frac{k+a-2}{2})!}{(\frac{a+k+2}{2})!(\frac{k+a-4}{2})!(\frac{k+a}{2})!(\frac{k+a-2}{2})!}\alpha^{\frac{k+a-2}{2}} + \cdots + \frac{(k-2)!(a+1)!-(k-3)!(a+2)!}{(k-2)!(a+1)!(k-3)!(a+2)!}\alpha^{a+2} + \frac{(k-1)!(a)!-(k-2)!(a+1)!}{(k-1)!(a)!(k-2)!(a+1)!}\alpha^{a+1}].$$

The geometric summation part of the above expression can be decomposed into a geometric summation of all the terms in the even case and an additional middle term of the odd-numbered series. Thereby,

$$(7) = W'(t)W^{k-1+a}(t) \times [\sum_{x=0}^{\frac{k-a-4}{2}}\frac{(k-2-x)!(a+1+x)!-(k-1-x)!(a+x)!}{(k-1-x)!(a+x)!(k-2-x)!(a+x+1)!} \times (\alpha^{k-1-x}-\alpha^{a+1+x})] + \frac{(\frac{k+a}{2})!(\frac{k+a-2}{2})!-(\frac{k-2+a}{2})!(\frac{k+a}{2})!}{(\frac{k+a}{2})!(\frac{k+a-2}{2})!(\frac{k-2+a}{2})!(\frac{k+a}{2})!}\alpha^{\frac{k+a}{2}}$$
$$= W'(t)W^{k-1+a}(t) \times [\sum_{x=0}^{\frac{k-a-3}{2}}\frac{(k-2-x)!(a+1+x)!-(k-1-x)!(a+x)!}{(k-1-x)!(a+x)!(k-2-x)!(a+x+1)!} \times (\alpha^{k-1-x}-\alpha^{a+1+x})]$$
$$< 0 \text{ , since } 0 \le a \le k-2 \text{ and } \alpha > 1.$$

Thus, $\frac{d[\frac{\sum_{j=0}^{k-1}\frac{W^j(t)}{j!}}{\sum_{j=0}^{k-1}\frac{R^j(t)}{j!}}]}{dt}$ is always < 0, for $R(t) = \alpha W(t)$. Following Theorem 1, we conclude that correct responses are stochastically faster than incorrect ones in the proportional Gamma model. □

As intimated earlier, the Smith and Van Zandt (2000) result on the conditional means is imposed by Theorem 3.

Corollary 3.1. *By the hierarchy of stochastic dominance relationships (see Townsend, 1990), it follows from Theorem 3 that the Smith and Van Zandt (2000) result follows immediately.*

Proof. Obvious. □

Moreover, ordinary Poisson counting models and their paired Gamma processing time distributions must also obey the implications of Theorem 3.

Corollary 3.2. *Since race models constructed from ordinary Gamma distributions, with one rate for 'rights' and another for 'wrongs' and rate for 'rights' > rate for 'wrongs', are special cases of proportional rates, inhomogeneous Poisson models, it follows that 'rights' are stochastically faster than 'wrongs' in the sense manifested above.*

Proof. Obvious. □

In order to offer a slightly less prosaic example of a proportional rates, inhomogeneous Gamma race model let $R(t) = at$ and $W(t) = bt$ where a and b are rate coefficients and let $a > b > 0$. Evidently, $\alpha = \frac{a}{b} > 1$. Fig. 4a shows densities given $a = 4, 10$ and $b = 2$, which leads to $\alpha = 2, 5$, respectively. Fig. 4b shows ratios of hazard functions $r(t)$ for each α. And Fig. 4c illustrates the differences of conditioned cdfs: $P(T \le t|R)$ - $P(T \le t|W)$, revealing their predicted dominance relationships.

3. Do inhomogeneous Poisson models exist that can predict faster wrongs than rights?

We do not have a complete solution to this intriguing question. It would be pleasant to find an inhomogeneous, and non-proportional-rates Poisson pair of counters that, violating one or

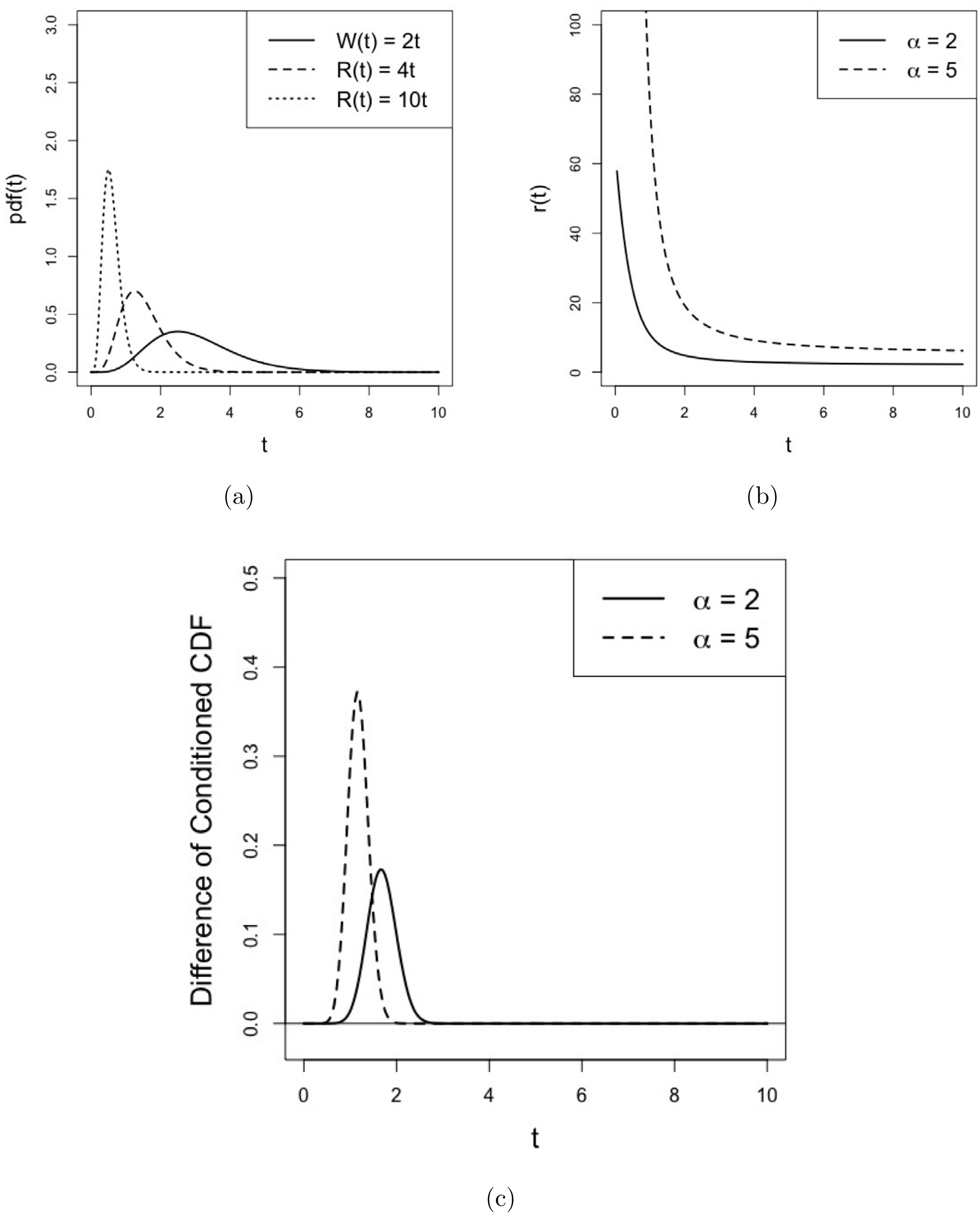

Fig. 4. (a) Densities of RTs for 'right' and 'wrong' channels, given $a = 4$, 10 and $b = 2$. (b) Decreasing ratios of hazard functions for $\alpha = 2$ ($a = 4$) and 5 ($a = 10$). (c) Differences of conditional cdfs $P(T \le t|R) - P(T \le t|W)$ are above 0 for $t > 0$, suggesting stochastically faster 'rights' than 'wrongs' for both α values.

more of the attendant axioms to Theorem 3, will always yield stochastically faster wrongs than rights. We so far have not found such a case.

However, the following example produced from the time-varying Poisson race model at least demonstrates the existence of a race of inhomogeneous Poisson counters which engender faster wrongs than rights for some values of t. In this example, the time-varying rate parameters of 'right' and 'wrong' are defined separately and are no longer proportional to each other. Results show that wrongs are stochastically faster than rights for the faster processing time but rights are faster for slower times.

Example. Let us still assume that the system follows a time-varying Poisson counting model with $R(t) = 7t+5$, $W(t) = 2t+4$, and k = 6. Following Eqs. (1) and (2), we have

$$g_W(t) = \frac{(2t+4)^5 2e^{-(2t+4)}}{5!}$$

$$g_R(t) = \frac{(7t+5)^5 7e^{-(7t+5)}}{5!}$$

and

$$\overline{G_W}(t) = \sum_{j=0}^{5} \frac{(2t+4)^j}{j!} e^{-(2t+4)}$$

$$\overline{G_R}(t) = \sum_{j=0}^{5} \frac{(7t+5)^j}{j!} e^{-(7t+5)}$$

We used Wolfram Mathematica to approximate the probability of being correct or wrong for this example. The probability of being correct, $P(R) = \int_0^\infty g_R(t)\overline{G_W}(t)dt = \frac{(1527194572016669)}{(80647217731665)e^3} \approx 0.9428$. The probability of being wrong, $P(W) = \int_0^\infty g_W(t)\overline{G_R}(t)dt = \frac{(144158031169111)}{(179216039403700)e^3} \approx 0.04005$.

The correspondent ratio of hazard (Eq. (3)) shown in Fig. 5a is non-monotonic (it increases with small t and then decreases with large t).

Now, we compare the processing time cumulative distributions conditional on correct and wrong responses, that is $P(T \leq t|R) = \frac{\int_0^t g_R(t')\overline{G_W}(t')dt'}{P(R)}$ vs. $P(T \leq t|W) = \frac{\int_0^t g_W(t')\overline{G_R}(t')dt'}{P(W)}$ at different values of t. As indicated in Fig. 5b, the difference between the cumulative processing time distribution conditional on correct responses is less than that of wrong response at small t, proving that correct responses are slower than incorrect ones with small t. Yet conditional cumulative processing time distribution of correct responses is larger than that of incorrect ones with larger t, showing that correct responses are faster than incorrect ones for large t. These observations are consistent with our theoretical predictions following Theorems 2 and 3.

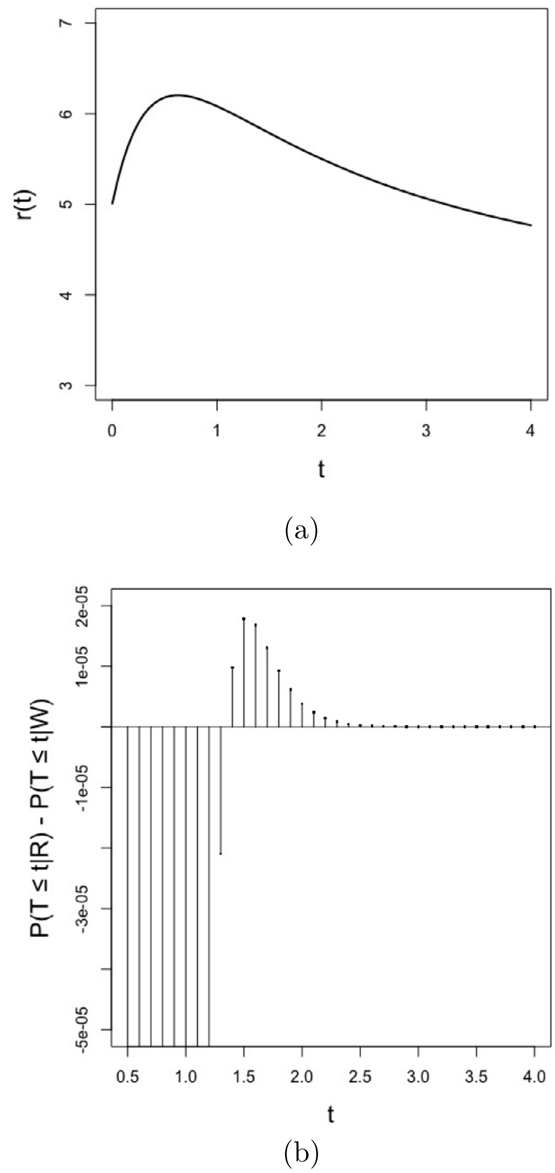

Fig. 5. (a) Non-monotonic ratio of 'right' hazard function over 'wrong' hazard function. (b) Difference between the cumulative processing time distributions conditional on 'rights' vs. 'wrongs'. With small t_i, differences are negative indicating stochastically faster 'wrongs' than 'rights'; whereas with large t_i, differences are above zero suggesting stochastically faster 'rights' than 'wrongs'.

4. Discussion

Race models can be classified as descriptive, meaning that the processing time distributions are defined directly on finishing times and there is no processing state–space for the channels. Contrarily, state–space race models assume each channel possesses a state–space which demarcates incremental accumulation of evidence for its associated decision and consequent response. For instance, Smith and Ratcliff (2009) developed a *dual diffusion model* that jointed these two aspects together. This model follows a race structure for decision process but involves separate Ornstein–Uhlenbeck diffusion process for each decisional accumulator.

Probability distributions for processing times can be compared via their means, medians or other more powerful characteristics such as order of their cumulative distribution functions, hazard functions and so on (Townsend, 1990). We thereby used such notions to beneficial effect in the present inquiry. We assumed the hazard function for 'rights' to be always greater than that for 'wrongs', in the class of descriptive race models. That is, $h_R(t) > h_W(t)$. This rule implies that their respective cumulative distribution functions are ordered as are their means and medians. We, and apparently many other cognitive-process modelers have harbored the belief that because independent race models have to predict that the likelihood of a 'right' response must be larger than 0.5, that 'rights' must be inherently faster than 'wrongs'. This intuition is in line with the rigorous proof by Smith and Van Zandt (2000) that for the class of inhomogeneous Poisson counter models, as long as the rates are proportional, the expected processing time for 'rights' is indeed quicker than that for 'wrongs'.

But we emphasized early on, that this does *not* imply that 'right' decisions (with time conditional on being 'right') are necessarily faster than 'wrong' decisions (time conditional on being 'wrong'). That is, the almost ubiquitous intuition mentioned above is mistaken. This realization led us to explore the actual distributions conditional on being 'right' vs being 'wrong'. Our current investigation demonstrates that the canonical intuition is incorrect if applied to the larger classes, firstly of descriptive race models, defined only in terms of their hazard functions and secondly of inhomogeneous Poisson counter models which do not necessarily obey the proportional rates dictum.

As proved in Theorem 1, a broad class of race models can produce either faster 'rights' or faster 'wrongs', depending on the monotonicity of their hazard function ratios. We examined these findings with a class of time-varying Poisson counting models in Theorems 2 and 3 and furthermore illustrated an example where faster 'wrongs' can be actually produced, at least at early times. Finally, with regard to this venture, the important question arises as to whether there exist inhomogeneous Poisson counting race models which induce 'wrongs' that are inevitably faster than 'rights' at our strong stochastic level. So far, we have not discovered such a model but we suspect they exist. We did uncover an example which for small (fast) times, does elicit faster 'wrongs' but instead produces faster 'rights' than 'wrongs' for long times. We hope our readers may solve this open problem and thereby discover inhomogeneous Poisson race models where the 'wrongs' are always faster than 'rights'.

In a slightly different direction, it may be intriguing to learn whether a state–space-based accrual halting model which employs monotonic increasing accruals, like our general Poisson counters, or has trajectories which can rise and fall before absorption, makes a fundamental difference in matters like our ratio of hazard functions, or not. It would evidently be useful to have and hand in hand analogous (to our results) basic, canonical results for other classes of models, such as more general diffusion models, the *linear ballistic accumulation* model (Brown & Heathcote, 2008) and so on. The only such characteristic of which we are aware, is the classic prediction of the Wiener diffusion and its close associates (e.g., the sequential probability ratio test model of Wald, 1947) that rights and wrongs are equally fast. An associated challenge for those models relying on probability mixtures of rates and/or criteria, is how to experimentally falsify these assumptions as opposed to simply enjoying their extra complexity in a more opportunistic way.

We view the questions posed in the present inquiry as concerning fundamental issues in human cognition. They complement other recent theoretical investigations into parallel processing systems. For instance, consider standard parallel models vs. standard serial models. Standard parallel models assume independent parallel processing durations while standard serial models assume independent successive durations and identically distributed times as well. Standard serial models can predict position effects through differing distributions on the processing order. The simplest standard parallel models have identically distributed channel processing times. However, this assumption is often abrogated in order to allow parallel models to predict position effects (e.g., Townsend & Ashby, 1983; Van Zandt & Townsend, 1993).

We recently demonstrated that standard parallel models will typically predict stochastically increasing intercompletion times and standard serial systems will typically predict positively correlated *total completion times* (total completion time is defined for each item by the time duration between the beginning of processing and the instant of completion of that specific item, whatever the architecture). Nonetheless, rather intriguingly, it turns out that both systems are controlled by their individual hazard functions (on channels, elements, stages, etc.) in the sense that certain dramatic alterations in these hazard functions can evoke the opposite kinds of behavior (Zhang, Liu, & Townsend, 2018; Zhang et al., 2019).

Other contemporary discoveries include foundational properties of mutually dependent (rather than independent) channels in parallel process models and their predictions for methodologically powerful workload capacity functions (Algom & Fitousi, 2016; Townsend, Liu, Zhang, & Wenger, 2020; Townsend & Wenger, 2004). There is a corresponding line of effort on parallel systems which are engaged in pattern classification types of computation. This research has been focusing primarily on their interactions (e.g., Musslick, Cohen, & Shenhav, 2019) and potential degradation thereby. We are hopeful that these up-to-now rather disparate tracks of research may begin to cross-fertilize each other.

Other future challenges, naturally, lie in the application to real-data response time distributions. One application is to use the ratio of hazard functions $r(t)$ for model selection. If a race model is chosen to account for empirically observed fast 'rights' or faster 'wrongs', then our current results ensure that the ratio of hazard functions produced by that model must obey the predicted monotonicity and therefore can rule out those that disagree with the predicted patterns.

With regard to the empirical side of matters, it will be important to learn first of all, if conditional dominance even holds at the level of the conditional distribution functions. After all, there was little at all known about stochastic dominance at different levels of power in response time studies before our explorations of these matters (Townsend, 1990; Townsend & Ashby, 1978; Townsend & Nozawa, 1995). It turned out, rather to our surprise, that ordinary kinds of data evinced rather remarkably strong levels of dominance. Whether this turns out to be true for conditional distributions as entertained here is an open question. Even more dramatic would be a finding that ratios of hazard functions could be monotonic, at least over reasonable durations of response times. Hazard functions themselves are non-trivial to estimate and their ratios as functions of time may be even more challenging.

References

Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of response times: Identification of cognitive mechanisms through mathematical modeling. In *The Oxford handbook of computational and mathematical psychology* (pp. 63–98). Oxford: Oxford University Press.

Algom, D., & Fitousi, D. (2016). Half a century of research on garner interference and the separability–integrality distinction. *Psychological Bulletin*, *142*(12), 1352.

Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, *44*(2), 310–329.

Audley, R. (1973). Some observations on theories of choice reaction time: Tutorial review. In *Attention and performance IV* (pp. 509–545). Academic Press New York.

Balakrishnan, J. (1994). Simple additivity of stochastic psychological processes: Tests and measures. *Psychometrika*, *59*(2), 217–240.

Bamber, D. (1969). Reaction times and error rates for 'same' - 'different' judgments of multidimensional stimull. *Perception & Psychophysics*, *6*(3), 169–174.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, *34*(3), 253–275.

Colonius, H., & Townsend, J. T. (1997). *Activation-state representation of models for the redundant-signals-effect*. New Jersey: Lawrence Erlbaum Associates Publishers.

Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology*, *38*(1), 35–58.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *34*(2), 187–202.

Ditterich, J. (2006a). Evidence for time-variant decision making. *European Journal of Neuroscience*, *24*(12), 3628–3641.

Ditterich, J. (2006b). Stochastic models of decisions about motion direction: behavior and physiology. *Neural Networks*, *19*(8), 981–1012.

Dunn, J. C., & Kalish, M. L. (2018). *State-trace analysis*. Berlin: Springer.

Hawkins, G. E., Mittner, M., Forstmann, B. U., & Heathcote, A. (2019). Modeling distracted performance. *Cognitive Psychology*, *112*, 48–80.

Heathcote, A., & Love, J. (2012). Linear deterministic accumulator models of simple choice. *Frontiers in Psychology*, *3*, 292.

Houpt, J. W., Townsend, J. T., & Jefferson, B. (2017). Stochastic foundations of elementary mental architectures. In *New Handbook of Mathematical Psychology*: *Vol. 2*.

Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, *121*(1), 1.

Khodadadi, A., & Townsend, J. T. (2015). On mimicry among sequential sampling models. *Journal of Mathematical Psychology*, *68*, 37–48.

Laming, D. R. J. (1973). *Mathematical psychology*. Cambridge, MA: Academic Press.

Lappin, J. S., & Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology*, *72*(6), 805.

Link, S., & Heath, R. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*(1), 77–105.

Luce, R. D. (1986). *Response times: their role in inferring elementary mental organization*. (8), Oxford: Oxford University Press.

Marley, A. (1971). Conditions for the representation of absolute judgment and pair comparison isosensitivity curves by cumulative distributions. *Journal of Mathematical Psychology*, *8*(4), 554–590.

Marley, A. A. J., & Colonius, H. (1992). The "horse race" random utility model for choice probabilities and reaction times, and its compering risks interpretation. *Journal of Mathematical Psychology*, *36*(1), 1–20.

McGill, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, & E. Galanter (Eds.), *Handbook of mathematical psychology*. New Your: Wiley.

McGill, W. J., & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, *2*(1), 1–18.

Miller, J. (1978). Multidimensional same–different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(3), 411.

Musslick, S., Cohen, J. D., & Shenhav, A. (2019). Decomposing individual differences in cognitive control: A model-based approach. In *Proceedings of the 41st annual meeting of the cognitive science society* (pp. 2427–2433).

Raab, D. H. (1962). Division of psychology: Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *245*(Series II), 574–590.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.

Smith, P. L. (2020). (2020). Time varying diffusion models, I. Time pressure, urgency, collapsing boundaries, and optimality. unpublished.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review*, *116*(2), 283.

Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, *53*(2), 293–315.

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, *32*(2), 135–168.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*(3), 251–260.

Swensson, R. G. (1972). The elusive tradeoff: Speed vs accuracy in visual discrimination tasks. *Perception & Psychophysics*, *12*(1), 16–32.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 168–199.

Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, *108*(3), 551.

Townsend, J. T., & Ashby, F. (1978). Methods of modeling capacity in simple processing systems. Hillsdale: Erlbaum.

Townsend, J. T., & Ashby, F. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: CUP Archive.

Townsend, J. T., & Diederich, A. (1996). The relationship of variance to interaction contrast in parallel systems factorial technology. *British Journal of Mathematical and Statistical Psychology*, *49*(2), 211–223.

Townsend, J. T., & Honey, C. J. (2007). Consequences of base time for redundant signals experiments. *Journal of Mathematical Psychology*, *51*(4), 242–265.

Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology*, *56*(6), 476–494.

Townsend, J. T., Liu, Y., Zhang, R., & Wenger, M. J. (2020). Interactive parallel models: no Virginia, violation of Miller's race inequality does not imply coactivation and yes Virginia, context invariance is testable. *The Quantitative Methods for Psychology*, *16*(2), 192–212, http://www.tqmp.org/RegularArticles/vol16-2/p192/p192.pdf.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359.

Townsend, J. T., & Wenger, M. (1996). Evidence monitoring theory: A dynamic extension of general recognition theory and cognitive stochastic processing theory. *Journal of Mathematical Psychology*, *40*(4), 353.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychological Review*, *111*(4), 1003.

Townsend, J. T., Wenger, M. J., & Houpt, J. W. (2018). Uncovering mental architecture and related mechanisms in elementary human perception, cognition, and action. In *Stevens' handbook of experimental psychology and cognitive neuroscience*: *Vol. 5*, (pp. 1–30). Wiley Online Library.

Van Zandt, T., & Townsend, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual and memory search: An evaluative review. *Perception & Psychophysics*, *53*(5), 563–580.

Wagenmakers, E. -J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*(1), 140–159.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica. Journal of the Econometric Society*, 279–313.

Wenger, M. J., & Townsend, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *The Journal of General Psychology*, *127*(1), 67–99.

Zhang, R., Liu, Y., & Townsend, J. T. (2018). A theoretical study of process dependence for standard two-process serial models and standard two-process parallel models. In *Invariances in human information processing* (pp. 117–142). Abingdon: Routledge.

Zhang, R., Liu, Y., & Townsend, J. T. (2019). A theoretical study of process dependence for critical statistics in standard serial models and standard parallel models. *Journal of Mathematical Psychology*, Article 102277.