Consequences of base time for redundant signals experiments

James T. Townsend*, Christopher J. Honey

Department of Psychology, Indiana University, Bloomington, IN 47405, USA

Abstract

We report analytical and computational investigations into the effects of base time on the diagnosticity of two popular theoretical tools in the redundant signals literature: (1) the race model inequality and (2) the capacity coefficient. We show analytically and without distributional assumptions that the presence of base time decreases the sensitivity of both of these measures to model violations. We further use simulations to investigate the statistical power model selection tools based on the race model inequality, both with and without base time. Base time decreases statistical power, and biases the race model test toward conservatism. The magnitude of this biasing effect increases as we increase the proportion of total reaction time variance contributed by base time. We marshal empirical evidence to suggest that the proportion of reaction time variance contributed by base time is relatively small, and that the effects of base time on the diagnosticity of our model-selection tools are therefore likely to be minor. However, uncertainty remains concerning the magnitude and even the definition of base time. Experimentalists should continue to be alert to situations in which base time may contribute a large proportion of the total reaction time variance.

1. Introduction

The incipient experimental psychology of the nineteenth century began to unravel the mechanisms responsible for sensory, cognitive and motor behavior. A great many of the strategies used *response times* along with various experimental conditions intended to assay the contributions of the psychological processes of interest in that period. A redoubtable challenge was then and is now, how to segregate the times of concern—for instance, an arithmetical computational mechanism—from those durations coming before or after. Early sensory preparation comes before, motor preparation and execution come after. Both of these contribute to what is known as the *base time* component of a reaction time.

Certain methodologies—such as Donders' (1869) *method of subtraction*, Sternberg's (1969) *additive factors method* and the later generalizations (Schweickert, 1978; Townsend & Ashby, 1983; Townsend & Schweickert, 1989)—have

witnessed some success in bypassing this obstacle. Nonetheless, the presence of the *base time* component of a response time continues to impede progress. Thus, investigators rarely fit response time distributions that account for a realistic probability distribution of the base time. In most cases, a constant is simply employed as the base time; sometimes it is entirely omitted.

Another approach is to call into play transform techniques (e.g. Fourier or Laplace transforms) that can be used, in principle, to segregate the base time from the processing distribution under study (Goldstone, 2000; Kohfeld, Santee, & Wallace, 1981). However, these strategies are prone to grave difficulties and risk (Sheu & Ratcliff, 1995; Smith, 1990). It is also known that the form of the base time distribution has consequences for parameter-fitting in numerous choice models (Ratcliff & Smith, 2004; Ratcliff & Tuerlinckx, 2002). We cannot review these rather vast subjects much less solve all the problems.

Our goal here is to investigate the *manner* and *magnitude* of the contamination by base time of two central constructs in elementary cognitive process theory and methodology: the race model inequality (Miller, 1982) and the capacity

---
*Corresponding author.

E-mail addresses: jtownsen@indiana.edu (J.T. Townsend), choney@indiana.edu (C.J. Honey).
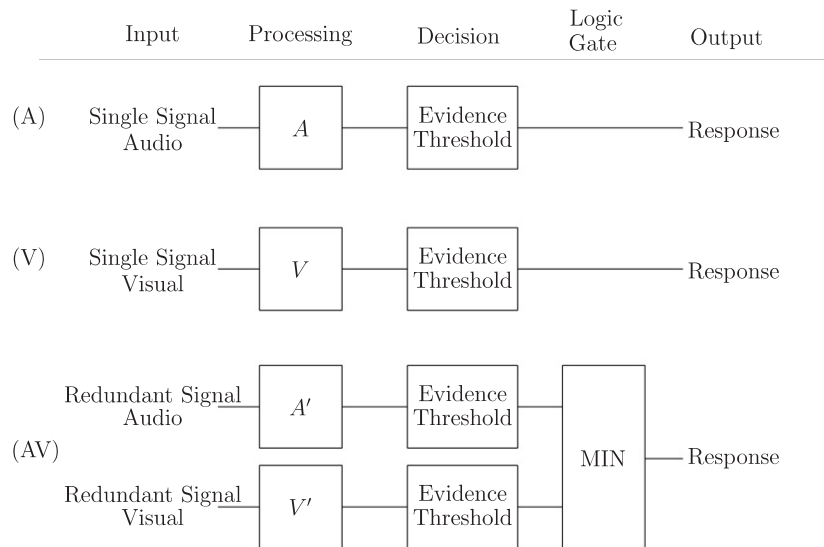
Fig. 1. Schematic of a standard redundant signals model.

coefficient (Townsend & Nozawa, 1995). Our investigation is latched to a certain, but very popular, paradigm centered on response times: the redundant signals design (RSD).

2. The RSD

In a RSD experiment we compare the reaction times to target stimuli presented singly with the reaction times to both target stimuli presented in combination. Consider, for the sake of concreteness, an audiovisual redundant signals experiment.[1] In the simplest case we will have three conditions: (1) presentation of an audio stimulus, (2) presentation of a visual stimulus and (3) presentation of audio and visual stimuli at the same time. Subjects are instructed to respond as soon as they detect an audio target or a visual target or both at the same time. In other words they are instructed to respond as soon as they detect a signal in either modality, regardless of what is occurring in the other modality. Trials in the first two conditions are called *single target* (SS) trials; trials in the third condition are called *redundant signal* (RS) trials.

There are two primary types of RSDs; the classification is based on what is presented in the *irrelevant modality* on SS trials. In the first type of RSD, the null stimulus is truly *null*; so on a single-signal trial in which the target appears in the visual modality, the experimenter might present *silence* on the (irrelevant) auditory modality. In the second type of RSD, the experimenter substitutes a *distractor* stimulus for the target. In this design it would not be silence, but rather a *non-target sound* presented in the (irrelevant) auditory modality when the target is visual.

In both kinds of RSD, and for a wide variety of experimental designs, it has frequently been observed that

reaction times are stochastically faster on the RS trials than on the single signal (SS) trials (Raab, 1962; Mordkoff & Miller, 1993; Westendorf & Blake, 1988). In other words, subjects are faster to respond to a flash and a beep presented together than they are to respond to the flash or beep presented alone. We label this kind of result—in which a significant speed superiority is found in moving from a single target condition to a multiple target condition—as a *redundant signals effect*.

2.1. Models explaining the RSs effect

In the search for an explanation of this phenomenon, it is natural to begin by positing some kind of interaction between the audio and visual modalities. Yet it has long been known that if the internal representation of the stimulus dimensions (features, etc.) is probabilistic, then RS facilitation can occur via statistical considerations alone (see e.g. Raab, 1962). We now formulate a framework in which to explicate the main aspects and questions relating to RS effects (see Fig. 1):

(1) We assume that the different modalities are processed in separate *channels*, along which evidence is accumulated toward the completion of a particular signal detection task. The audio channel accumulates evidence toward detection of audio targets; the visual channel accumulates information toward detection of visual targets. We further assume that the signals impinging on the channels are processed in *parallel*: one channel does not have to wait for the other channel to finish before it can begin processing its input (but see Townsend & Nozawa, 1997).

(2) We assume that the rate of processing along each channel is invariant across the SS and RS conditions (the audio signal is processed in the same way regardless of whether a visual signal is also presented

---

[1]Nothing that follows hinges on this particular choice. The results presented here could apply to any experimental design in which a subject monitors multiple sources of information for a rapid response.

on that trial, and vice versa). This assumption, known as *context invariance*,[2] is crucial, because it is the theoretical link that justifies our comparison of data gathered in the SS condition against data gathered in the RS conditions.

(3) We assume that on the RS trials there is a parallel *race* between the audio and the visual processing channels. Successful completion of the signal detection process on *either* channel is sufficient to produce a motor response, and so the reaction time measured on an RS trial will simply be equal to the process completion time of the *faster* channel on that trial. This assumption is also known as the assumption of a *minimum-time stopping rule*, since the completion time of the entire system is the minimum of the completion times of the individual channels.

(4) We sometimes assume that the signals impinging on each channel are processed *independently*: the rate at which the audio signal is processed is not affected by the rate at which the visual signal is processed on that trial, and vice versa.

A model that satisfies assumptions 1 through 3 is commonly referred to as a *race model*. A model that satisfies assumptions 1 through 4 is a particular type of race model; it is known, in terms of the taxonomy of information-processing models (Townsend & Ashby, 1983; Townsend & Nozawa, 1995), as an Unlimited Capacity, Independent, Parallel (UCIP) model with a minimum-time stopping rule. In this paper, when we say *race model* we mean any model satisfying assumptions 1 through 3. When we say *UCIP model* we mean any model satisfying assumptions 1 through 4, i.e. we mean a race model with independent channels.[3]

We model the cognitive processing times on each channel using a non-negative random variable; times on the audio channel are modeled with the variable A and times on the visual channel are modeled with the variable V. These variables are theoretical quantities. The variables governing the empirical *reaction time* in each condition are written as RT_A and RT_V, respectively. When base time is neglected in the model, then we assume that each measured reaction time is a sample realization of the corresponding theoretical quantity:

$$
\begin{aligned}
\text{Audio SS trial:} & \quad RT_A = A, \\
\text{Visual SS trial:} & \quad RT_V = V, \\
\text{Audiovisual RS trial:} & \quad RT_{AV} = \min(A', V'),
\end{aligned}
\tag{1}
$$

where the random variables $A =_{df} A'$ and $V =_{df} V'$. The notation $X =_{df} Y$ means that "X and Y have the same

distribution functions". See Appendix D for a complete list of notations and abbreviations.

Now any race model (including the UCIP model) can account for the RSs effect at the level of *mean* reaction times; the observed speedup in the RS condition is a simple consequence of the fact that for any two random variables, $E[\min(A, V)] \leqslant \min(E[A], E[V])$ from which it follows that we will experimentally observe $E[RT_{AV}] \leqslant \min(E[RT_A], E[RT_V])$. However, of all race models, the UCIP model provides the most parsimonious explanation since it does not require any channel interactions and yet still produces a RSs effect.

A major step forward occurred when Jeffrey Miller (Miller, 1978, 1982) pointed out that all race models had to satisfy a simple yet powerful inequality. At the time, context invariance was not made explicit as a necessary assumption; this theoretical gap was filled independently by Luce (1986) and Ashby and Townsend (1986). Slightly modifying Miller's original development to make this assumption more evident we see the argument as follows.

Each reaction time we record is understood to be a sample from one of the random variables $RT_A = A$, $RT_V = V$ or $RT_{AV} = \min(A', V')$. From these data we can estimate the cumulative distribution functions $F_{RT_A} = P(RT_A \leqslant t)$, $F_{RT_V} = P(RT_V \leqslant t)$ and $F_{RT_{AV}} = P(RT_{AV} \leqslant t)$. Miller (1982) showed that, for any race model,

$$
\begin{aligned}
F_{RT_{AV}}(t) &= F_{\min(A',V')}(t) \\
&= F_{\min(A,V)}(t) \quad (CI) \\
&= P(\min(A, V) \leqslant t) \\
&= P(A \leqslant t) + P(V \leqslant t) - P(A \leqslant t \text{ and } V \leqslant t) \\
&= F_A(t) + F_V(t) - P(A \leqslant t \text{ and } V \leqslant t) \\
&= F_{RT_A}(t) + F_{RT_V}(t) - P(A \leqslant t \text{ and } V \leqslant t). \tag{2}
\end{aligned}
$$

Notice that we used *context invariance* to obtain the line labeled (CI). Context invariance is the assumption that the channel process variables are invariant across SS and RS trials, i.e. $A' =_{df} A$ and $V' =_{df} V$. The assumption of context invariance provides a theoretical bridge between our RS and our SS data. Following from the final line above, we simply notice that we are subtracting a positive quantity from the right-hand side, so that

$$
F_{RT_{AV}}(t) \leqslant F_{RT_A}(t) + F_{RT_V}(t). \tag{3}
$$

We emphasize that this inequality holds for any race model, not only for the independent (UCIP) case. Eq. (3) is known variously as "Miller's Race Model Inequality" or simply as the "Race Model Inequality" and has been widely used in psychophysics to test for interactions between posited perceptual processes. Maris and Maris (2003) have recently developed a non-parametric test which provides a principled means of determining the level of statistical significance of any violations observed empirically. Note that if the inequality is violated we falsify not only the UCIP model, but in fact *any* race model of task performance, since the race model inequality is expected to

---

hold regardless of the dependence between A and V. In the event that violations are detected, we can test alternative hypotheses that posit one or another sort of *coactivation* between process channels (Miller, 2004; Schwarz & Ischebeck, 1994; Townsend & Wenger, 2004).

Bear in mind, however, that the race model inequality *can hold* in the absence of independence, and also that it *can fail to hold* for a race model in which one or other assumptions are violated. In real-time functioning systems, dependencies can readily alter the marginal processing time distributions, thus contravening context invariance. In fact, it was shown (Townsend & Wenger, 2004) that mutually inhibitory or mutually facilitatory systems very often produce, in the first case, marginal cumulative distributions less than the single target distributions and in the second case, marginal cumulative distributions greater than the single target distributions.

The race model inequality is a conservative test (Patching & Quinlan, 2004; Schwarz & Ischebeck, 1994) but it has been, and continues to be, widely employed in the analysis of reaction time data. It is therefore important that we understand the implications of base time for the diagnosticity of this test.

3. Base time

3.1. The what and where of base time

In the explication above we assumed that our measured reaction times RT_A, RT_V and RT_{AV} were samples of the process variables A, V and $\min(A, V)$. But it is clear that any observed reaction time will comprise not only the completion times of the cognitive processes in which we are interested, but also additional components such as (a) the time taken for the basic sensory system to transmit to higher processing centers and (b) the time taken to prepare and execute the motor response. The time taken to perform these other functions is commonly called *base time*. [4]

Our special concern in this paper is the theoretical and practical impact of base time in race models of reaction time. For instance, we want to know what happens to the validity of Eq. (3) when the empirically measured RT_A is not a sample from the random variable A but is rather a sample from $A + B_A$ where B_A is a random variable modeling base time. We make the canonical assumption that base time is *invariant* across all experimental conditions (SS-Audio, SS-Video and RS) and also *independent* of our channel process variables A and V (Luce, 1986; Ulrich & Giray, 1986). These assumptions are a kind of *context invariance* for base time; if base time is really a proxy for a functionally distinct motor response time then they are plausible. There are, however, some significant methodological and ontological concerns with these

---

[4]It is also referred to as "residual time" (e.g. Luce, 1986) and "non-decisional time" (e.g. Ratcliff & Tuerlinckx, 2002).

assumptions (Dzhafarov, 1992), and we return to this issue in Section 5.1.

Following Ulrich and Giray (1986) and with reference to Fig. 2, our race model incorporating base time is

Audio SS trial	$RT_A = A + B_A,$
Visual SS trial	$RT_V = V + B_V,$
Audio visual RS trial	$RT_{AV} = \min(A', V') + B_{AV},$

$$(1^*)$$

where $A =_{df} A'$, $V =_{df} V'$ and $B_A =_{df} B_V =_{df} B_{AV}$. Notice that we allow for the possibility of a dependence between A' and V', so this is a general race model. Also notice that here, and for the rest of this paper, we assume that the distribution of the base time is invariant across conditions so that all base time variables follow share an identical density function, $f_B(t)$. We will further assume the notation that the variables X, X', X'', etc. are identically and independently distributed. These variables are introduced in order to make clear that although the variables across conditions are equal in distribution (usually due to context invariance), their realizations are entirely independent. See Appendix A for some clarifications and caveats concerning the use of random variable notation.

For the model just described, the following chain of inequalities is provided in the literature:

$$F_{\min(A', V') + B_{AV}}(t) \tag{4a}$$

$$\leqslant F_{\min(A + B_A, V + B_V)}(t) \tag{4b}$$

$$\leqslant F_{A + B_A}(t) + F_{V + B_V}(t). \tag{4c}$$

The first inequality, (4a) $\leqslant$ (4b), was stated by Ulrich and Giray (1986) under the assumption that all the random variables are independent; it was later proven by Colonius (1990) in the more general case in which dependence between A and V is allowed. The second inequality, (4b) $\leqslant$ (4c) is just the race model inequality, Eq. (3), with some relabeling of variables.

Suppose now that we were to simulate a UCIP model with a base time component, i.e. we were to simulate the system depicted in Fig. 2 with the stipulation that variables A and V are independent. Then following from Eq. (4), which is expressed in terms of theoretical process variables, we expect the following properties to emerge in our simulation data:

$$F_{RT_{AV}}(t) \tag{5a}$$

$$\leqslant F_{\min(RT_A, RT_V)}(t) \tag{5b}$$

$$\leqslant F_{RT_A}(t) + F_{RT_V}(t). \tag{5c}$$

We see that the inequality (5a) $\leqslant$ (5c) is identical to the race model inequality (3); the only difference is that the quantities in Eq. (5) incorporate base time while those in Eq. (3) do not. Clearly, then, violations of the race model inequality falsify race models whether or not the race models include base time (Ulrich & Giray, 1986). What is not so evident, and what we demonstrate in Section 3.2, is that the presence of base time serves to *decrease* the maximum sensitivity of the race model inequality (and related) tests to the detection of race model

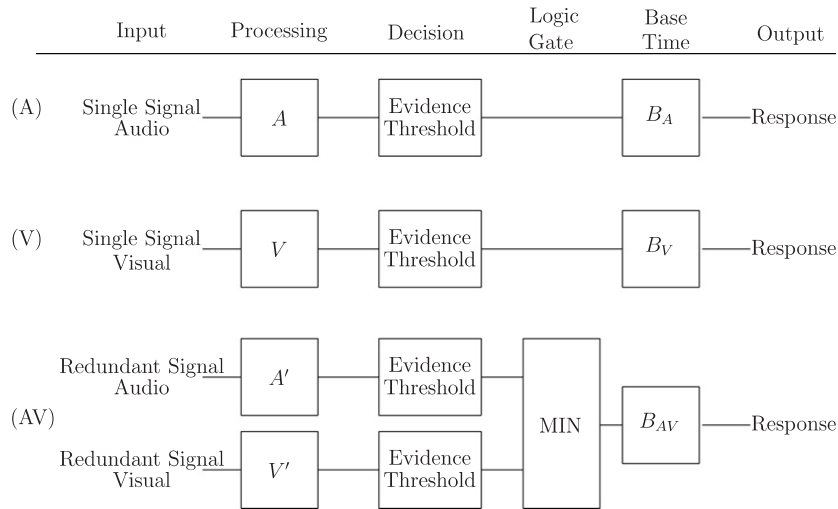|	Input	Processing	Decision	Logic Gate	Base Time	Output
(A)	Single Signal Audio	A	Evidence Threshold		B_A	Response
(V)	Single Signal Visual	V	Evidence Threshold		B_V	Response
(AV)	Redundant Signal Audio	A'	Evidence Threshold	MIN	B_{AV}	Response
	Redundant Signal Visual	V'	Evidence Threshold			

Fig. 2. Schematic of a redundant signals model including base time as in Eq. (1*).

violations. More precisely: if we begin with a system that does not include base time but does contain a dependence between A and V such that the system produces violations of the race model inequality, *then* the incorporation into that system of base time components B_A, B_V and B_{AV} can only serve to reduce the magnitude of the exhibited violations or to mask them entirely.

We emphasize that it is the *ideal* (infinite sample) sensitivity that is impacted in this way by the introduction of base time. The effects on test diagnosticity are more complicated when only small data samples are available; some exploratory results concerning power and false alarm rates are provided in Section 3.4.

3.2. Base time as a filter

Given any two independent random variables, say X and B, whose probability density functions $f_X(t)$ and $f_B(t)$ are known, we can calculate the density function, $f_{X+B}(t)$, of the variable $X + B$ by convolution:

$$f_{X+B}(t) = \int_{-\infty}^{\infty} f_X(s)f_B(t-s)\,ds = f_X * f_B(t).$$

When both X and B are non-negative variables this equation can be rewritten as

$$f_{X+B}(t) = \int_{-\infty}^{t} f_X(s)f_B(t-s)\,ds = f_X * f_B(t)$$

and since random variables modeling process times are non-negative we will assume this simpler convolution form for the remainder of this paper.

It is also well known (e.g. Townsend & Ashby, 1983) that we can obtain an analogous result for the cumulative distribution functions:

$$\underbrace{F_{X+B}(t)}_{\text{Output}} = \int_{-\infty}^{t} F_X(s)f_B(t-s)\,ds = \underbrace{F_X}_{\text{Input}} * \underbrace{f_B(t)}_{\text{Kernel}}.$$

When an input function is convolved with another function (especially a unimodal function with smaller support) then it is usual to refer to the simpler function as a *kernel* and to view the *output* of the convolution transformation as a filtered version of the *input*. We say that $F_X(t)$ is being *filtered* by $f_B(t)$, and the terminology is appropriate because f_B, the probability density function of our base time variable, is likely to have a small half-width and to be unimodal.[5] Kernels of this type can be easily understood to act as local averaging mechanisms: each point of the output function is a kernel-weighted average of corresponding points in the input function. We now proceed to apply this reasoning to the quantities involved in our RSs paradigm.

When we test the race model inequality, Eq. (3), we are interested in the roots of the *race model test function*

$$R^*(t) \equiv F_{AV}(t) - F_A(t) - F_V(t), \tag{6}$$

where A, V and AV are the "true" cognitive process times in the SS-audio, SS-video and RS conditions, respectively. If $R^*(t) > 0$ for any time t then the race model inequality is violated.[6] Unfortunately we do not observe $R^*(t)$. Instead we observe the empirical quantity

$$R_B^*(t) \equiv F_{RT_{AV}}(t) - F_{RT_A}(t) - F_{RT_V}(t). \tag{7}$$

---

[5] If base time comprises mostly motor response time then we have reason to believe this (Luce, 1986; Meijers & Eijkman, 1974). Ratcliff and Tuerlinckx (2002) make a similar assumption when they calculate the effects of base time on the estimation of the parameters of diffusion models—see their Figure 11.

[6] Here we follow the notation of Colonius and Diederich(2006) who analyze the closely related race function:

$$R = F_{RT_{AV}}(t) - \min[F_{RT_A}(t) + F_{RT_V}(t), 1].$$

Unfortunately the convolution analysis we present here does not appear to have direct applicability to their approach because of this definitional difference.

There is, however, a simple way of relating the theoretical $R^*(t)$ and the empirical $R_B^*(t)$:

$$R_B^*(t) \equiv F_{RT_{AV}}(t) - F_{RT_A}(t) - F_{RT_V}(t)$$
$$= F_{AV+B_{AV}}(t) - F_{A+B_A}(t) - F_{V+B_V}(t)$$
$$= \int_{-\infty}^{t} F_{AV}(s) f_B(t-s)\, ds$$
$$- \int_{-\infty}^{t} F_A(s) f_B(t-s)\, ds - \int_{-\infty}^{t} F_V(s) f_B(t-s)\, ds$$
$$= \int_{-\infty}^{t} [F_{AV}(s) - F_A(s) - F_V(s)] f_B(t-s)\, ds$$
$$= \int_{-\infty}^{t} R^*(s) f_B(t-s)\, ds.$$
$$= R^* * f_B(t). \tag{8}$$

The linearity of the convolution operator and the assumption that all base time variables are described by a common density function $f_B(t)$, provide us with the result that observed $R_B^*(t)$ is a filtered version of the "true" quantity $R^*(t)$ upon whose properties we would like to base our model selection decisions.

By definition we know that $R^*(0) = 0$ and that asymptotically we have $R^*(\infty) = -1$. We want to know whether there is some positive time t^* for which $R^*(t^*) > 0$ because the existence of such a time would constitute a violation of the race model inequality.[7]

We can use a standard result from the theory of function convolution (Kecs, 1982) to demonstrate that the maximum positive value of $R^*(t)$ is always greater than the maximum positive value R_B^*. More generally, using Hölder's Inequality, it can be shown that when we filter an input function with a probability density function the output function will always have a smaller absolute magnitude at its extrema than the input function. The main theorem that we need is from the theory of convolution operators (Kecs, 1982).

Lemma 1. *If $f \in \mathscr{L}^1(\mathbb{R}^n)$ and g is bounded on $\mathbb{R}$ then the convolution $f * g$ exists, is bounded, continuous and*

$$\|f * g\|_{L^\infty} \leqslant \|f\|_{L^1} \|g\|_{L^\infty}, \tag{9}$$

where $\|f\|_{L^p}$ denotes the Lipschitz p-norm of the function f, and where $\mathscr{L}^1(\mathbb{R}^n)$ denotes the set of Lebesgue-integrable functions on $\mathbb{R}^n$.

Proof. See Corollary 2 in Kecs (1982, p. 66). □

This lemma is in fact a special case of a more general theorem that provides a bound on the L_p-norm of a convolution product, with the bound depending on the L_p-norms of the two functions being convolved.

Now we apply this bound to our purpose. First, we choose $f = f_B(t)$ since the density function $f_B \in \mathscr{L}^1(\mathbb{R}^n)$ with the additional property that $\int_{-\infty}^{\infty} |f_B(t)|\, dt =$

$\|f_B\|_{L^1} = 1$. Secondly, we choose $g = R^*(t)$ since $R^*(t)$ is bounded. Then, recalling that $\|g\|_{L^\infty} = \max_t |g(t)|$ and substituting in Eq. (9) we obtain

$$\max_t |R_B^*(t)| = \|R^* * f_B(t)\|_{L^\infty}$$
$$\leqslant \|f_B\|_{L^1} \|R^*\|_{L^\infty}$$
$$= \|R^*\|_{L^\infty}$$
$$= \max_t |R^*(t)|.$$

This result does not yet get us what we want, because, as a result of the fact $R^*(\infty) = -1$, we will always have $\max |R^*(t)| = 1$. This makes the bound too loose to significantly constrain the height of any region of positivity, which is usually much smaller than 1. However, the previous derivation demonstrates the general principle that a local maximum of a bounded function will be filtered into a smaller local maximum when the kernel is a probability density function. This makes sense, since we expect a local averaging process to flatten any peaks in the input function. It will take a little more work in order to express this mathematically.

Suppose that the function $R^*(t) = g(t)$ is indeed positive over some range of t before returning to its asymptote at $g(\infty) = -1$. We want to know what the filtering process does to the height of this 'hump' in g. In order to do so, we first notice that we can decompose any function g into positive and negative pieces

$$g^+(t) = \begin{array}{ll} g(t) & \text{if } g(t) > 0, \\ 0 & \text{otherwise,} \end{array}$$

and

$$g^-(t) = \begin{array}{ll} g(t) & \text{if } g(t) < 0, \\ 0 & \text{otherwise,} \end{array}$$

such that $g(t) = g^+(t) + g^-(t)$. We will need the following simple result:

Lemma 2. *Let $f, g \in L^1(\mathbb{R})$. Let $f(t) \geqslant 0 \, \forall t \in \mathbb{R}$. Then*

$$(f * g)^+ \leqslant f * g^+$$

on the entire domain $\mathbb{R}$.

Proof.

$$f * g = \int_{-\infty}^{t} f(s) g(t-s)\, ds$$
$$= \int_{-\infty}^{t} f(s) g^+(t-s)\, ds + \int_{-\infty}^{t} f(s) g^-(t-s)\, ds$$
$$\leqslant \int_{-\infty}^{t} f(s) g^+(t-s)\, ds \quad \text{since } f(s) \geqslant 0. \tag{10}$$

Therefore

$$[f * g]^+ \leqslant \left[\int_{-\infty}^{t} f(s)g^+(t-s)\,ds \right]^+$$

$$= \int_{-\infty}^{t} f(s)g^+(t-s)\,ds \quad \text{again since } f(s) \geqslant 0$$

$$= f * g^+. \quad \square \tag{11}$$

Now if the kernel $f(t) \geqslant 0$ for all t, as is the case for any density function that models reaction times, then we have

$$\max_t \{(f * g)^+\} \leqslant \max_t \{f * g^+\} \tag{12a}$$

$$\leqslant \max_t \{g^+\}, \tag{12b}$$

where the first line follows from Lemma 2 and the second line follows from our convolution magnitude bound, Lemma 1. This can all be brought together as follows:

Theorem 3. *When a processing system contains a base time component that can be modeled additively (i.e. via convolution of existing cdf's with a density function), this component tends to decrease the maximal height of any region of positivity in our empirically obtained race model test function,* R_B^*. *In other words,* $\max_t\{(R^* * f_B)^+\} \leqslant \max_t\{(R^*)^+\}$. *The presence of base time thus tends to decrease the magnitude (i.e. maximum height) of any observed race model violations.*[8]

This result informs us about the maximum height of the positive component of the race model test function. We now move on to consider the *area* under the positive component of the race model test function. In doing so, the first result we employ is a standard relationship from the theory of convolution, relating the integral of two functions and the integral of their convolution:

Lemma 4. *Let* $f, g \in L^1(\mathbb{R}^n)$. *Then* $f * g$ *exists almost everywhere and*

$$\int_{-\infty}^{\infty} f * g(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx \int_{-\infty}^{\infty} g(x)\,dx.$$

Proof. This result follows quite directly from the definition of convolution and an application of Fubini's Theorem. A proof can be found as Proposition 3 in Kecs (1982, p. 62). $\square$

As before, we decompose g into positive and negative components, and then the results from Lemmas 2 and 4 can be applied to show that the *area* of a region of positivity (i.e. the integral of g^+) in the output function is always less than the area of the corresponding regions in the input function:

$$\int_{-\infty}^{\infty} (f * g)^+(x)\,dx$$

$$\leqslant \int_{-\infty}^{\infty} f(x) * g^+(x)\,dx \quad \text{from Lemma 2}$$

$$= \int_{-\infty}^{\infty} f(x)\,dx \int_{-\infty}^{\infty} g^+(x)\,dx \quad \text{from Lemma 4}$$

$$= \int_{-\infty}^{\infty} g^+(x)\,dx. \tag{13}$$

This last result can be summarised as:

Theorem 5. *When a processing system contains a base time component that can be modeled additively (i.e. via convolution of existing cdf's with a density function), this component tends to decrease the total area of positivity in our empirically obtained race model test function,* R_B^*. *In other words,* $\int_{-\infty}^{\infty}(R^* * f_B)^+\,dt \leqslant \int_{-\infty}^{\infty}(R^*)^+\,dt$. *The presence of base time thus tends to decrease the magnitude (i.e. area) of observed race model violations.*

Theorems 3 and 5 are our central analytical results about the impact of base time on model-selection tools based on additive relations between reaction time CDF's. These results solidify our intuitions about what happens when we add a base time 'filter' to our statistical signature, the race function R^*. See Figs. 4 and 5 for examples of these effects. Typically, the base time filter smoothes functions and squashes them toward the ordinate axis.

Consider, at one extreme, the Dirac Delta function. Filtering with $\delta(t-a)$ does not alter the shape of the original function at all, but simply produces an output that is phase shifted a units. At the other extreme we might filter using a density function which is a constant height h over an interval of width $1/h$; filtering with this function produces, for each time T, the arithmetic mean of the input function over the range $[T-h, T]$. If h is small then the output will be a strongly smoothed version of the input.[9]

As a rule of thumb,[10] the flatter and wider the kernel function, the more extreme the smoothing effect that is produced by convolution. If base time has a large variance then its probability density function might well be a flat, wide kernel. This means that violations of the race model inequality will be more difficult to detect in data the greater the variance of the base time. In fact, it is possible for the presence of base time to have a large enough effect that a theoretical $R^*(t)$ exhibiting race model violations shows up empirically as an observed $R_B^*(t)$ which does not. In other

---

[8]It is possible that our observed function R_B^* contains more than one positive local maximum. If this were to occur then the result would imply that the largest of *all* the maxima in the output function, R_B^*, would still be smaller than the largest of *all* the maxima in the input function, R^*.

[9]Note that filtering using the Dirac Delta function corresponds to the situation in which base time has zero variance so that $f_B(t) = \delta(t-\mu)$. Filtering with a constant function corresponds to the base time variable following a uniform distribution.

[10]In general the effects can be quite subtle and will depend on the shapes of all the functions involved, but here we limit ourselves to relatively smooth and symmetrical functions.

words, it is sometimes possible to entirely 'filter away' a region of positivity in $R^*(t)$.

3.3. Quantitative estimates of smoothing effects

To provide some more intuition, Figs. 3, 4 and 5 give an idea of what filtering effect sizes might look like in a standard experimental setting. We assume the form of realistic "true" cognitive process time distributions that fix the value of $R^*(t)$ as in Eq. (6) and we further assume a realistic base time distribution so that we can calculate $R_B^*(t)$ as in Eq. (7). Any difference between $R^*(t)$ and $R_B^*(t)$ must then be due to contamination of our data by base time, and the magnitude of the difference informs us of the magnitude of the contamination effect. Specifically, we chose

Processing time on redundant trials : AV is distributed as N(330, 40),

Processing time on audio SS trial : A is distributed as N(410, 40),

Processing time on visual SS trial : V is distributed as $Exp(\frac{1}{500})$,

Base time components $B_A =_{df} B_V =_{df} B_{AV}$ are distributed as N(150, σ).

These distributions were chosen so that the means and variances would be somewhat realistic for a RSs experiment, and also so that the function R^* would exhibit a change of sign; the particulars of our distributional choices are less important than the fact that the R^* and R_B^* curves would appear typical to an experimentalist. Modeling process times using normal distributions are unrealistic
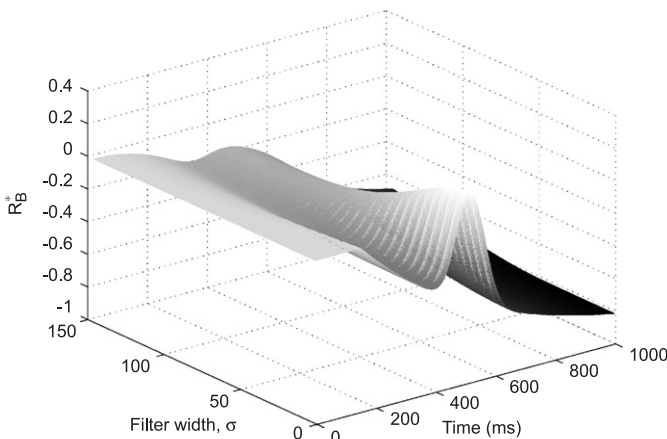

Fig. 3. The output function, $R_B^* = R^* * f_B$ shown as a function of the common standard deviation of the base time components B_A, B_V and B_{AV}. When $\sigma = 0$ the function $f * g$ exhibits a region of positivity near $t = 500$ ms. For larger σ the function $f * g$ is negative definite; the region of positivity has filtered away. Notice how the smoothing effect of the convolution operation increases as we increase σ.
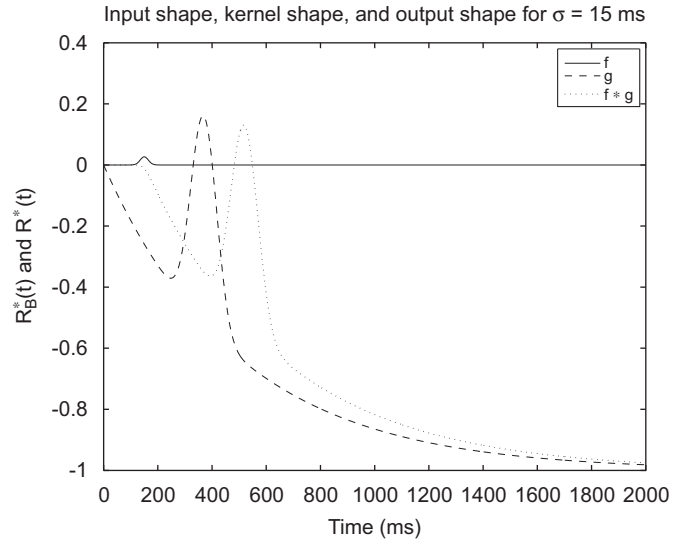

Fig. 4. The kernel g is a Normal density function with mean $\mu = 150$ ms and standard deviation $\sigma = 15$ ms; we notice that convolution with kernel g does not mask the positivity in the original function f.
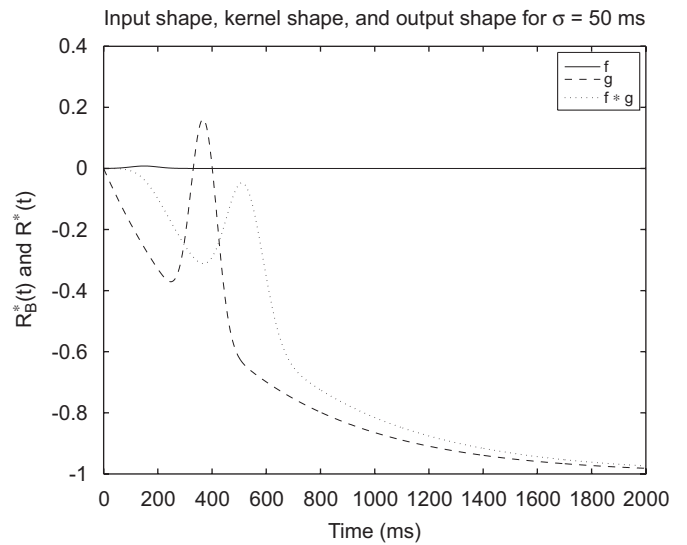

Fig. 5. The kernel g is a Normal density function with mean $\mu = 150$ ms and standard deviation $\sigma = 50$ ms; we notice that convolution with kernel g does indeed mask the positivity in the original function f.

insofar as it allows for negative times. However, even for the highest variance parameter choices we use, negative reaction times occur on fewer than 0.1% of trials.

In Fig. 3 we illustrate the effects of varying σ, the standard deviation of the base time variable. Figs. 4 and 5 provide more detail, showing the input function, the kernel and the output function for two particular values of σ. In Fig. 4 the kernel is a Normal density function with mean $\mu = 150$ ms and standard deviation $\sigma = 15$ ms; the filtering process does not mask the positivity in the original function. In Fig. 5 we see the same input function, but now the filter is a Normal density function with $\mu = 150$ ms and $\sigma = 50$ ms; this wider, flatter kernel does indeed mask
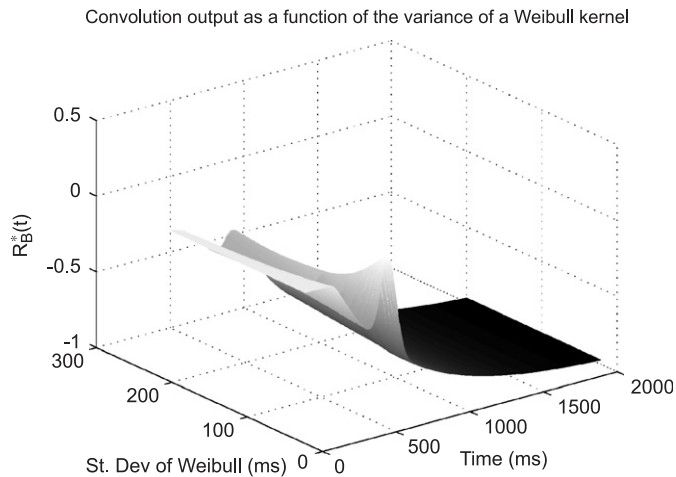
Fig. 6. The output function, $R_B^* = R^* * f_B$ shown as a function of the standard deviation of the base time variable. The base time is Weibull distributed with scale parameter 200 and a varying shape parameter. The smoothing effect of the convolution operation increases as we increase σ in a very similar manner to the effect produced by the Gaussian kernel.

the positivity in the original function.[11] In Fig. 3 we see the convolution output, $R_B^* = R^* * f_B$, plotted against the standard deviation, σ, of the filter.

We also performed simulations to examine the smoothing effects of uniform and Weibull distributions, while retaining the same distributions on channels A, V and AV as described above. For both Weibull and uniformly distributed base time kernels the results were very similar to what was obtained using the Gaussian kernel. Fig. 6, for example, shows the effects of smoothing using a Weibull-shaped kernel with scale parameter 200 ms and a varying shape parameter. The results are plotted as a function of the standard deviation of the Weibull kernel for easier comparison with Fig. 3. The equivalence across distribution types is suggestive that, for the parameter ranges relevant to RSDs, the smoothing effects depend only weakly on the functional form of the base time distribution. This is unsurprising since, as pointed out by Ratcliff and Tuerlinckx (2002), the variance of the base time is typically much smaller than the variance of the processing times; this implies that the kernel will only have a 'local' smoothing effect, and so its precise shape is less important.

It is apparent that wider filters do indeed produce larger smoothing effects. It is also apparent that asymmetric kernels can produce different deformations of the input functions, but these effects will depend on the shapes of both the kernel and the function that it is deforming. The

distributions for $f_B(t)$ that we implement here are in accordance with estimates obtained experimentally by (Kohfeld et al., 1981) and also informally in our laboratory. However, Luce (1986) points out that asymmetric, long-tailed estimates of $f_B(t)$ have been reported (Snodgrass, 1969).

3.4. Considerations of statistical power

The analyses above show how model selection based on the quantity $R_B^*(t)$ (which is contaminated by base time) will differ from model selection based on the uncontaminated quantity $R^*(t)$. However, those analyses were conducted under the assumption that we had in our possession perfect estimates of the cumulative distribution functions required to calculate $R_B^*(t)$. This assumption will never be satisfied in practice, and so here we investigate how empirical model selection is affected by sample size, base time variance and other parameters. Although a thorough investigation of the statistical power of race model tests (i.e. the probability of rejecting the race model given that the data were not generated by a race model) is beyond the scope of this paper, we present some preliminary comments based on the results of Monte Carlo simulations.

For the simulations that we report, the processing times on channels A and V were modeled as Gaussians with means fixed at 400 ms and standard deviations fixed at 100 ms. Following Ratcliff and Tuerlinckx (2002) we model the base time as a uniform distribution whose range was varied. For each set of parameters, reaction time data was randomly generated, a sample was taken from the generated data and then a model selection was conducted based on the sampled cdf's. By repeating this process 1000 times for each parameter setting we obtained estimates of the statistical power (correct rejection rate) as well as the correct retention rate. Three parameters were varied in this exploratory analysis:

- N, the sample size. This is the number of reaction times that were sampled in each condition (A, V, and AV) and used to estimate the cdf's F_{RT_A}, $F_{RT_V}(t)$ and $F_{RT_{AV}}$, respectively. Simulations were performed with $N = 100, 250$ and 500 samples.
- s_B, the base time variance. The base time was drawn from a uniform distribution of mean 250 msec whose range was set at either 50, 100 or 200 ms. These ranges correspond to standard deviations of 15, 29 and 58 ms, respectively.
- RSF, a parameter that governs the amount of RSs facilitation included when generating data. The RSF is defined as the difference between the mean value of the processing times on the AV channel and the mean value of the quantity $\min(A, V)$, which is what the mean processing time would be equal to under a race model. When RSF is zero we are simulating a race model. When RSF is positive (negative) the amount of facilitation in

---

[11]When considering whether a given filter can mask all traces of positivity in the original $R^*(t)$, it is worth considering the amount of negativity in $R^*(t)$ that is present *before* the t-value at which the positivity first emerges. The way that a filter masks a positive region is by averaging that positive region with a negative region. If these negative regions do not exist, then there is no negativity that can 'average out' the positive regions. If no negativity precedes a region of positivity, then that region positivity cannot be entirely masked. The worst that can happen is that its height and area will be attenuated by a region of negativity that *follows* it.
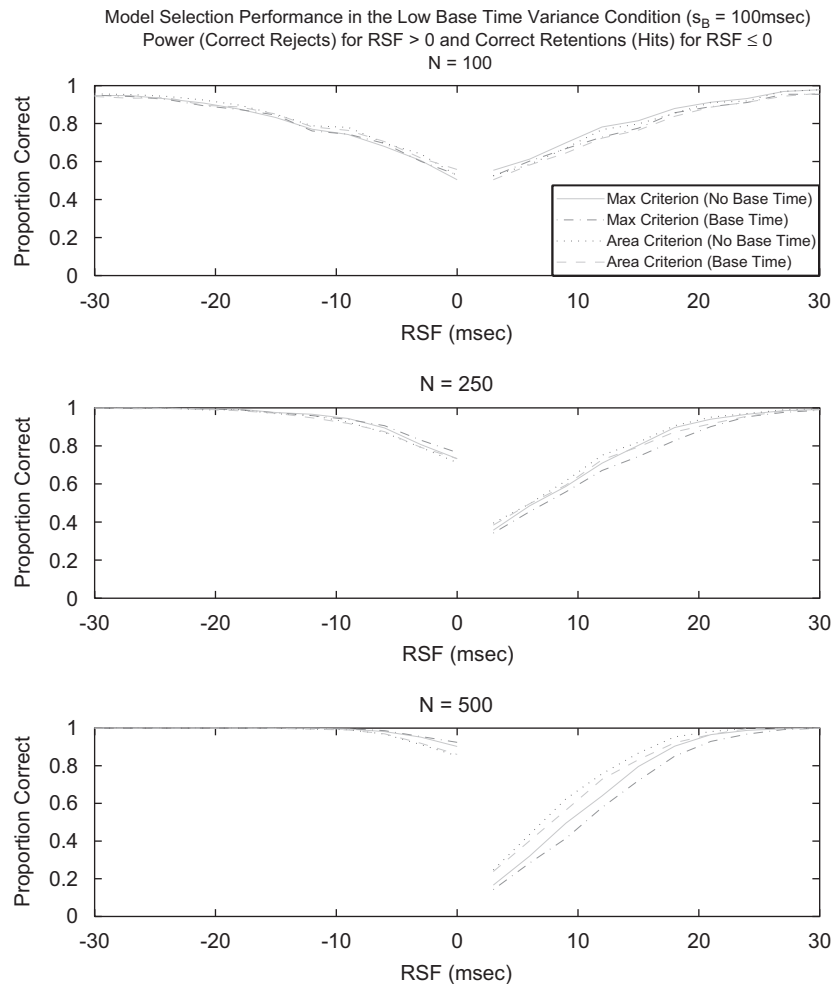
Fig. 7. The probability of making the correct model selection decision as a function of redundant signals facilitation in the generating model, and also as a function of the absence or presence of base time. Data shown here are for *low* variance base time, i.e. $s_B = 100$ ms. When $RSF < 0$ the generating model produces less facilitation than expected from a UCIP race model: the correct decision is to retain the null (race) model; a rejection of the race model constitutes a false alarm. When $RSF > 0$ the generating model produces more facilitation than expected from a UCIP race model, and the correct decision is to reject; the probability of these correct rejections constitutes the statistical power of the our model selection tool.

the RSs condition is greater (less) than what would be expected under an independent channels (UCIP) race model.

To calculate power it is necessary to set a criterion for rejection of the null hypothesis (i.e. for rejection of the race model). Here we used two different model selection criteria:

The 'Max' Criterion: We reject the null hypothesis if the maximum value of the empirical $R_B^*(t)$ exceeds a value of 0.05 at any time.

The 'Area' Criterion: We reject the null hypothesis if the total area underneath the positive component of the empirical $R_B^*(t)$ exceeds a threshold value of 2 ms.

The thresholds of 0.04 and 2 ms were chosen by trial and error to produce decent model selection performance.[12] Varying these thresholds produces the familiar signal detection trade-off in which greater statistical power is

accompanied by a higher rate of incorrect retentions of the race model. It is worth noting that if one chose to use a criterion in which the race model is rejected whenever $R_B^*(t)$ exhibits any positivity whatsoever (i.e. if one sets the 'max' threshold to 0) then one obtains extraordinarily high incorrect rejection rates even when RSF is negative, because the $R_B^*(t)$ function is always near zero for small values of t, and so minuscule fluctuations at early times inevitably produce small regions of positivity in the race model test function.

Figs. 7 and 8 show how the probability of making the correct model selection decision (retain when $RSF \leqslant 0$ and reject when $RSF > 0$) varies with RSF and with N in the medium variance ($s_B = 100$ ms) and high variance ($s_B = 200$ ms) conditions, respectively. Of course the results shown in these plots are contingent upon our choice of model rejection criteria under both 'Max' and 'Area' thresholds. A more general picture (in which no threshold is privileged) is presented in Figs. 9 and 10, which show how the maximum of $R_B^*(t)$ and the area under the positive

---

[12]To get a sense for the size of these thresholds, notice that the positive peak in Fig. 4 has a maximum value of about 0.16 and an area of 7.3 ms.
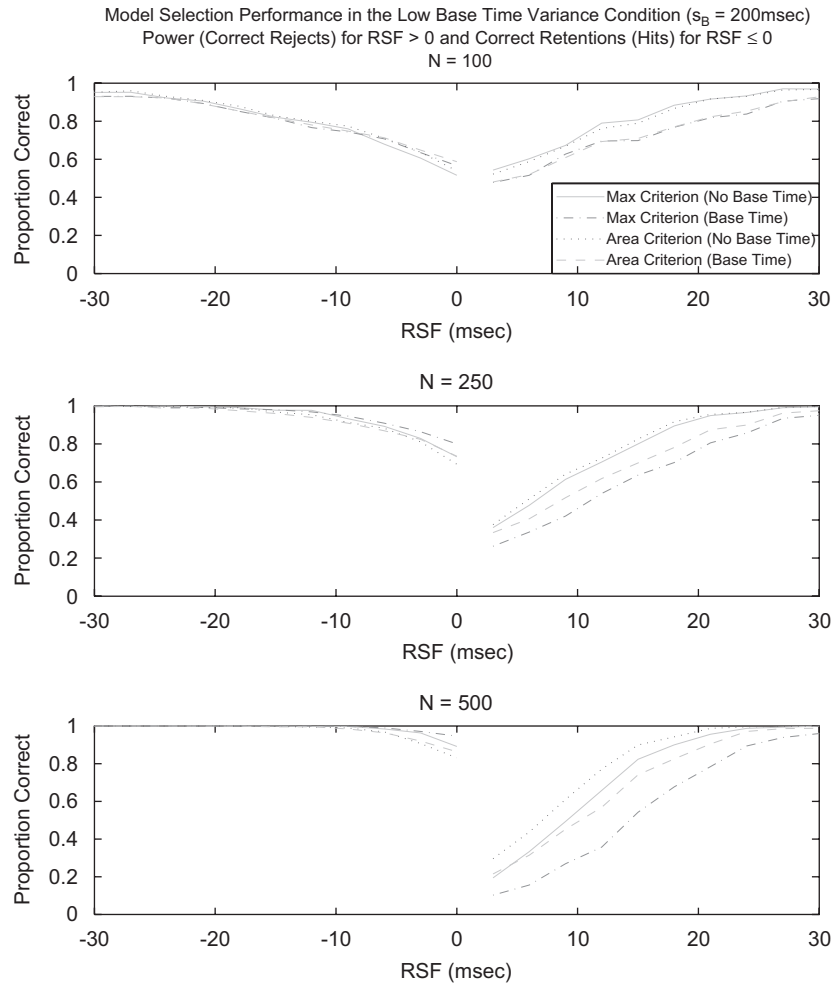
Fig. 8. The probability of making the correct model selection decision as a function of redundant signals facilitation in the generating model, and also as a function of the absence or presence of base time. Data shown here is for *high* variance base time, i.e. $s_B = 200$ ms. When $RSF < 0$ the generating model produces less facilitation than expected from a UCIP race model: the correct decision is to retain the null (race) model; a rejection of the race model constitutes a false alarm. When $RSF > 0$ the generating model produces more facilitation than expected from a UCIP race model, and the correct decision is to reject; the probability of these correct rejections constitute the statistical power of the our model selection tool.

component of $R_B^*(t)$ are *distributed* as we vary the amount of true RSs facilitation while fixing the parameters $N = 250$ and $s_B = 100$.

In interpreting these data, we draw two preliminary conclusions:

(1) The effects of base time on the distributions of both the maximum value and the positive area of the sampled $R_B^*(t)$ are small across a wide range of realistic choices of the base time range s_B. Power is decreased and the rate of incorrect rejections is increased by the presence of base time, especially for small values of RSF (i.e. for very difficult model selection decisions). However, the effects shown in the figures are produced for quite large values of the base time variance parameters: when the base time has a range of 50 ms about its mean then the plots of statistical power with and without base time are indistinguishable.

(2) Decreasing N tends to produce the same effect as decreasing the model rejection thresholds: it produces a

decrease in power and an increase in the rate of correct retentions. This trade-off is illustrated in Figs. 7 and 8. However, there is an exception to this overall pattern. We observe that, for small positive values of RSF, increasing N actually *decreases* power regardless of the absence or presence of base time. This phenomenon can be explained in terms of the fact that the race model inequality is, as mentioned earlier, a conservative test. So even if we have a perfect (infinite sample) estimate of the race model test function $R^*(t)$—a function *not* contaminated by base time—we are still led to incorrectly retain the null hypothesis when the RSF is small. Since tests based on the race model inequality are intrinsically biased toward falsely accepting the null hypothesis, we can actually increase our probability of correctly rejecting the null hypothesis by decreasing sample size. We note, however, that these effects are only present for small values of RSF. Figs. 7 and 8 attest that for $RSF > 25$ ms (i.e. even when there is only a relatively small speed superiority present over the
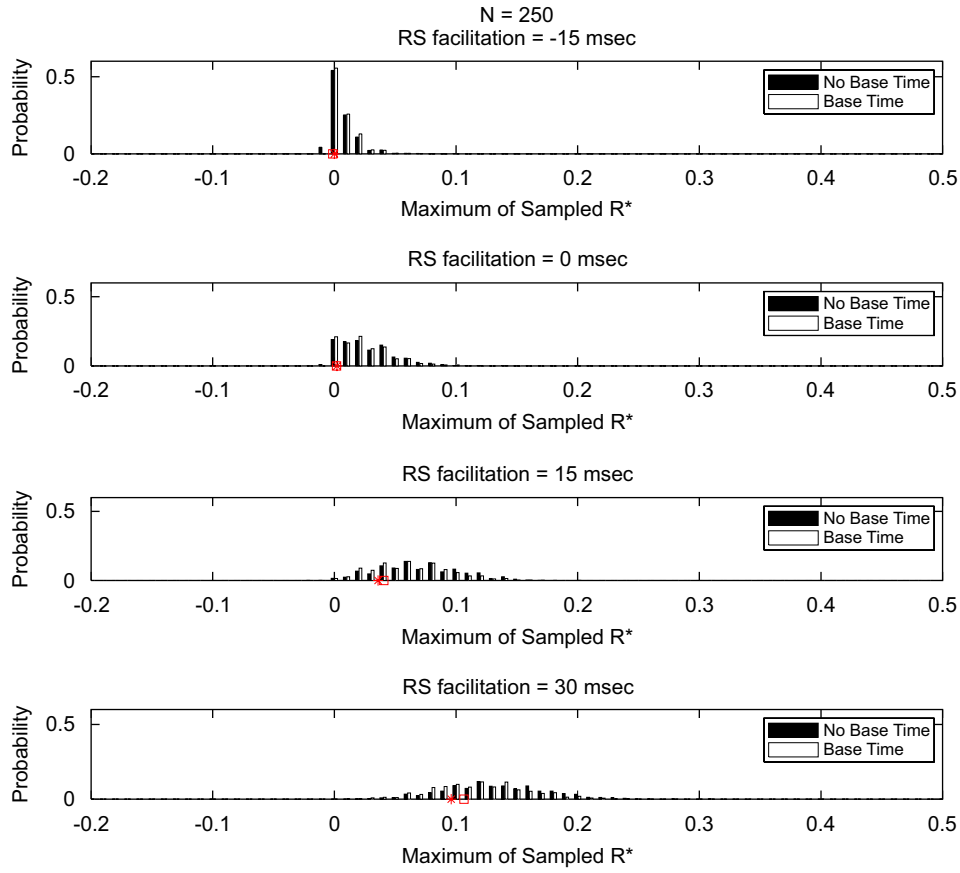
Fig. 9. The distribution of the *maximum value* of the sampled $R^*(t)$ (and $R_B^*(t)$) functions with and without base time. When base time is present, its range is set at $s_B = 100$ ms. The maximum values of the true (infinite sample) $R^*(t)$ and $R_B^*(t)$ functions are indicated by a square and an asterisk, respectively, on the *x*-axes of each plot.

UCIP model), statistical power is over 80% for all of the values of N and s_B that we considered and for both 'Max' and 'Area' threshold criteria.

The applicability of these conclusions is restricted to the class of unsophisticated (although commonly employed) threshold-based model selection heuristics we describe here. We would expect that the non-parametric race model test proposed by Maris and Maris (2003) could achieve much better model selection performance. We note, however, that their test is not designed for a pre- or post-decisional base time of the type we implement here, and, as the authors themselves point out, the power of their test is decreased when this form of base time is present in the processing system. It would be interesting to know how large this decrease in power is. A more systematic investigation of power should also investigate a wider range of distributions and parameterizations for both decisional and non-decisional variables, as well as different types of model rejection criteria. Finally, its seems that the presence of positivity in $R_B^*(t)$ is less diagnostic when it occurs at small values of t than when it occurs at large values of t, and we suggest that this phenomenon could be fruitfully investigated in future studies.

3.5. Relating the race model inequality and the capacity coefficient

Above we explore the effects of base time on the diagnosticity of the test provided by the distributional inequalities in Eq. (5). The reader might well be inclined to ask why it is that we persist in testing (5a) $\leqslant$ (5c) in our data when testing (5a)$\leqslant$(5b) appears to provide a tighter bound. The reason is that we do not have any general experimental procedure for estimating the quantity $F_{\min(RT_A,RT_V)}(t)$ which constitutes the tighter bound (5b).

Recall that $F_{\min(RT_A,RT_V)}(t) = F_{A+B_A}(t) + F_{V+B_V}(t) - F_{A+B_A \text{ and } V+B_V}(t)$. Our SS data provide us with estimates of $F_{A+B_A}(t)$ and $F_{V+B_V}(t)$. But we are denied access, in a RSs experiment, to the joint distribution $F_{A+B_A \text{ and } V+B_V}(t)$.

It is, however, possible to estimate the joint distribution $F_{A+B_A \text{ and } V+B_V}(t)$ in the case when there are *no dependencies* between channels A and V, because in that case

$$F_{A+B_A \text{ and } V+B_V}(t) = F_{A+B_A}(t)F_{V+B_V}(t), \qquad (14)$$

and the joint quantity on the RHS is a simple product of cdf's estimated from our audio and visual SS data. So if we are willing to assume independence of A and V (i.e. we are willing to assume a UCIP model) then we can indeed test
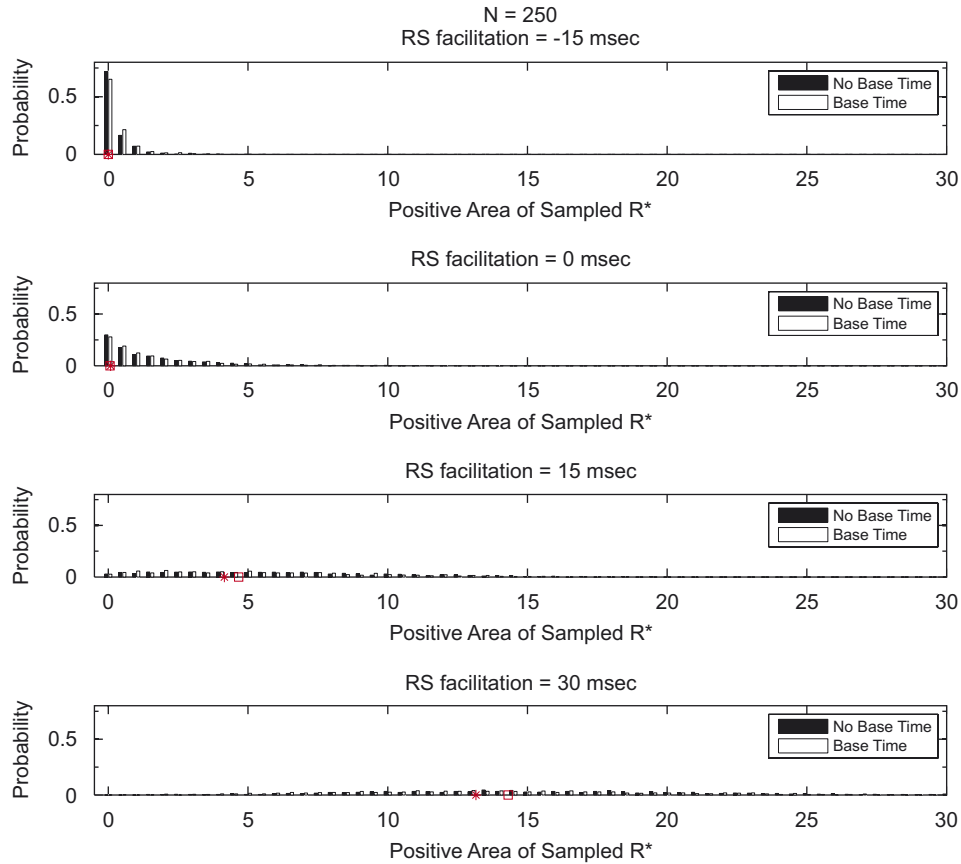
Fig. 10. The distribution of the *area* under the positive component of the sampled $R^*(t)$ (and $R_B^*(t)$) functions with and without base time. When base time is present, its range is set at $s_B = 100$ ms. The areas under the positive component of the true (infinite sample) $R^*(t)$ and $R_B^*(t)$ functions are indicated by a square and an asterisk, respectively, on the x-axes of each plot.

Table 1
The cdf-inequalities that can be practically employed used to test the race models and UCIP models both in the presence and the absence of base time

	Race model	UCIP model
Base time excluded	$F_{RT_{AV}} \leqslant F_{RT_A} + F_{RT_V}$	$F_{RT_{AV}} = F_{RT_A} + F_{RT_V} - F_{RT_A}F_{RT_V}$
Base time included	$F_{RT_{AV}} \leqslant F_{RT_A} + F_{RT_V}$	$F_{RT_{AV}} \leqslant F_{RT_A} + F_{RT_V} - F_{RT_A}F_{RT_V}$

our model using the inequality (5a)$\leqslant$(5b) rather than the standard (5a)$\leqslant$(5c). This is the essential reasoning from which we construct another benchmark tool for RSs architectures: the capacity coefficient $C(t)$ (Townsend & Nozawa, 1995; Wenger & Townsend, 2001, Faces as Gestalt Stimuli: Process Characteristics).

The distinction between model-selection based on the race model inequality (as discussed above) and model-selection based on the capacity coefficient (a quantity that we discuss presently in Section 4) is essentially this:

- the race model inequality is a test that benchmarks RSs facilitation relative to Eq. (5c) and which can falsify *all* race models;
- the capacity coefficient is benchmarked relative to Eq. (5b) and is therefore a more specific test that can be used to falsify *independent-channels* race models, i.e. UCIP models.

Of course we have not yet defined the capacity coefficient (see Section 4) but we wished to make clear the relationship between these two measures. It is also the case that the usefulness of the race model bound and the capacity coefficient is not restricted to formal model-selection; both measures can be used by experimentalists to benchmark the magnitude of the facilitation effects that they observe. In particular, the capacity coefficient has been related to a variety of assumptions about processing architecture, stopping rules and attentional allocation (e.g. Townsend & Nozawa, 1995; Townsend & Wenger, 2004). These results, in turn, allow predictions that latch system characteristics to the major distributional inequalities.

In order to further clarify the situation, Table 1 below summarizes the cdf-inequalities that can be practically employed to test race models and UCIP models, both with and without base time. The race model is commonly expressed as a difference of cdf's as in (7), while the

capacity coefficient is defined using logarithmic transforms of cdf's as in (15). However, the theoretical foundations linking the two measures are quite straightforward as we will soon demonstrate, and as summarized in Table 1.

For the race model, the bounds do not change when base time is included, but they are weakened for the reasons we have outlined above. For UCIP models, the test is weakened from equality to inequality; in Section 4.3 we discuss the possibility of formulating a UCIP model test that takes better account of base time. All distribution-level tests should optimally be performed within a principled statistical framework. (Maris & Maris, 2003, but see also our comments on this approach in the discussion.).

4. Measures of capacity

Capacity is a measure of how the processing rates on individual channels vary as we vary the number of active channels. In an attempt to nail down this slippery concept in the RSs context, Townsend and Nozawa (1995) defined $C(t)$, a non-negative real-valued index that is called the *capacity coefficient*. $C(t)$ can be understood as measuring how efficiency (the amount of total information processed per unit of time) varies with load (the amount of incoming information that is being processed). We know that the UCIP model predicts a specific increase in efficiency as load is increased, and we *normalize* the capacity coefficient relative to this specific amount so that the reaction time data gathered from a pure UCIP model will produce $C(t) = 1$. The investigator then compares empirical data against the benchmark $C(t) = 1$ to see whether subjects exhibit more or less RSs facilitation than that produced by a UCIP model.

The capacity coefficient and the Miller bound are transformations of the same data, and so it is no surprise that base time affects both of them. Townsend and Nozawa (1995) noted that the presence of base time should lead the capacity coefficient to underestimate capacity; more recently Ingvalson and Wenger (2005) demonstrated this fact using stochastic simulations. Here we present the first analytic results describing the relationship between base time and capacity. Before doing so, however, we recap some of the properties of $C(t)$ and their relationship to the results from the preceding section.

4.1. Capacity without base time

In the absence of base time, for the redundant targets experimental setup, capacity is defined as

$$
\begin{aligned}
C(t) &\equiv \frac{H_{AV}(t)}{H_A(t) + H_V(t)} \\
&= \frac{\log[1 - F_{RT_{AV}}(t)]}{\log[1 - F_{RT_A}(t)] + \log[1 - F_{RT_V}(t)]},
\end{aligned}
\tag{15}
$$

where $H_A(t), H_V(t)$ and $H_{AV}(t)$ are the integrated hazard functions calculated in the SS-Audio, SS-Visual and RS

conditions, respectively (Townsend & Nozawa, 1995). We mention the hazard functions in order to demonstrate their concrete ties to the quantities F_{RT_A}, F_{RT_V} and $F_{RT_{AV}}$ that were the focus of the preceding sections, but we emphasize that none of these quantities account for base time in their standard definitions.

In order to clarify the relationship between the performance of the capacity measure in the absence and presence of base time, we will compare two derivations. In the derivation that follows we show how a UCIP model that *excludes* base time produces $C(t) = 1$, and in Section 4.2 we show how the analogous derivation for a base time *inclusive* UCIP model produces $C(t) \leqslant 1$. So we begin by assuming a UCIP model without base time, i.e. we assume that processing on the channels is independent, that context invariance holds, and that processing is carried out as schematized in Fig. 1. In that case we can unpack $C(t)$ into

$$
\begin{aligned}
C^*(t) &= \frac{\log[1 - F_{\min(A',V')}(t)]}{\log[1 - F_A(t)] + \log[1 - F_V(t)]} \quad \text{(CI)} \\
&= \frac{\log[1 - F_{\min(A',V')}(t)]}{\log[(1 - F_A(t))(1 - F_V(t))]} \\
&= \frac{\log[1 - (F_{A'}(t) + F_{V'}(t) - F_{A' \text{ and } V'}(t))]}{\log[1 - F_A(t) - F_V(t) + F_A(t)F_V(t)]} \\
&= \frac{\log[J(t) + F_{A' \text{ and } V'}(t)]}{\log[J(t) + F_A(t)F_V(t)]} \\
&= \frac{\log[J(t) + F_{\max(A',V')}(t)]}{\log[J(t) + F_A(t)F_V(t)]},
\end{aligned}
\tag{16}
$$

where $-1 \leqslant J(t) \leqslant 1$ is simply some quantity that appears on both the numerator and denominator of our expression. On the line marked (CI) we have invoked context invariance, i.e. we assume that marginal cdf's on the channels are invariant across conditions so that $F_{RT_{AV}}(t) = F_{\min(A',V')}(t)$, where $A =_{df} A'$, $V =_{df} V'$.

Working from the last line in the sequence above we can finally invoke the independence assumption of the UCIP model, substituting $F_{A' \text{ and } V'}(t) = F_{\max(A',V')}(t) = F_A(t)F_V(t)$, and from this it immediately follows that

$$
C^*(t) = 1,
$$

a result which holds for any UCIP model without base time.

This derivation was sketched under UCIP assumptions, but we remind the reader that the capacity coefficient can still be interpreted when neither context invariance nor channel independence is satisfied. The capacity measure provides us with a general measure of processing efficiency relative to the UCIP baseline:

$C(t) > 1$. *Supercapacity*: The processing efficiency on individual channels *increases* as the number of active channels (i.e. the amount of information being fed into the system) is increased. (See Townsend & Nozawa, 1995 for more details.)

$C(t) = 1$. *Unlimited capacity*: The processing efficiency on individual channels is *invariant* as the number of active

channels (i.e. the amount of information being fed into the system) is increased.

$C(t) < 1$. *Limited capacity*: The processing efficiency on individual channels *decreases* as the number of active channels (i.e. the amount of information being fed into the system) is increased.

4.2. Capacity with base time

We now move on to consider whether and how our interpretations of $C(t)$ should be modified when we take base time into account. What happens to the capacity measurement of our benchmark UCIP model when base time is added to it? As in Section 4.1 we assume independence and context invariance, but now we substitute base-time-augmented cdf's into the capacity equation:

$$
\begin{aligned}
C_B^*(t) &= \frac{\log[1 - F_{\min(A', V') + B_{AV}}(t)]}{\log[1 - F_{A + B_A}(t)] + \log[1 - F_{V + B_V}(t)]} \\
&= \frac{\log[1 - F_{\min(A', V') + B_{AV}}(t)]}{\log[(1 - F_{A + B_A}(t))(1 - F_{V + B_V}(t))]} \\
&= \frac{\log[1 - \int_{-\infty}^{t}(F'_A(t - s) + F_{V'}(t - s) - F_{A' \text{ and } V'}(t - s))f_B(s)\, ds]}{\log[1 - F_{A + B_A}(t) - F_{V + B_V}(t) + F_{A + B_A}(t)F_{V + B_V}(t)]} \\
&= \frac{\log[1 - F_{A' + B_A}(t) - F_{V' + B_V}(t) + \int_{-\infty}^{t} F_{A' \text{ and } V'}(t - s)f_B(s)\, ds]}{\log[1 - F_{A + B_A}(t) - F_{V + B_V}(t) + F_{A + B_A}(t)F_{V + B_V}(t)]} \\
&= \frac{\log[J(t) + \int_{-\infty}^{t} F_{A' \text{ and } V'}(t - s)f_B(s)\, ds]}{\log[J(t) + F_{A + B_A}(t)F_{V + B_V}(t)]} \\
&= \frac{\log[J(t) + F_{\max(A', V') + B_{AV}}(t)]}{\log[J(t) + F_{A + B_A}(t)F_{V + B_V}(t)]},
\end{aligned}
\tag{17}
$$

where once again $-1 \leqslant J(t) \leqslant 1$ is some quantity common to both the numerator and denominator. It should not be difficult to see that this derivation is analogous to the one that was performed for the UCIP model without base time. But the value of the capacity coefficient now depends on the relative magnitude of $F_{\max(A', V') + B_{AV}}$ and $F_{A + B_A}(t)$ $F_{V + B_V}(t)$, whereas previously it depended on the relative magnitude of $F_{\max(A', V')}(t)$ and $F_A(t)F_V(t)$. In the previous derivation the quantities on numerator and denominator were *equal* if A and V were independent, but now that base time is incorporated this is no longer the case. Instead of equality, an *inequality* holds between the quantities in the numerator and the denominator. The following theorem demonstrates this fact.[13]

Lemma 6. If $A =_{df} A' =_{df} A''$, $V =_{df} V' =_{df} V''$ and $B_A =_{df}$ $B_V =_{df} B_{AV}$ are all mutually stochastically independent non-negative random variables then

$$
F_{\max(A', V') + B_{AV}}(t) \geqslant F_{A + B_A}(t)F_{V + B_V}(t).
\tag{18}
$$

Proof. We begin by noting the cdf-level equality

$$
{}_A(t) + F_V(t) = F_{\min(A'', V'')}(t) + F_{\max(A', V')}(t)
\tag{19}
$$

---

[13]The primed variables, e.g. A, A', A'', are used for the sake of accuracy and generality in the proof, but they do complicate the notation somewhat. For an initial understanding of the result it may be helpful simply to ignore the primes.

which holds regardless of the dependency between A and V.

Convolving both sides of Eq. (19) with the density function of some positive random variable B we obtain the relation

$$
\begin{aligned}
F_{A + B}(t) + F_{V + B'}(t) &= F_{\min(A'', V'') + B''}(t) \\
&\quad + F_{\max(A', V') + B'''}(t)
\end{aligned}
\tag{19$'$}
$$

between the marginal cdf's. We can rewrite this as

$$
\begin{aligned}
F_{A + B_A}(t) + F_{V + B_V}(t) &= F_{\min(A'', V'') + B_{AV}}(t) \\
&\quad + F_{\max(A', V') + B'_{AV}}(t)
\end{aligned}
\tag{19$''$}
$$

by choosing $B =_{df} B_A =_{df} B_V =_{df} B_{AV} =_{df} B'_{AV}$.

Now we recall that Colonius (1990) has shown that

$$
F_{\min(A'', V'') + B_{AV}}(t) \leqslant F_{\min(A'' + B_A, V'' + B_V)}(t)
\tag{20}
$$

and by substituting (19$''$) into (20) we obtain

$$
\begin{aligned}
F_{\max(A', V') + B_{AV}}(t) &\geqslant F_{A + B_A}(t) + F_{V + B_V}(t) \\
&\quad - F_{\min(A'' + B_A, V'' + B_V)}(t).
\end{aligned}
\tag{21}
$$

Now we use the fact that $A =_{df} A' =_{df} A''$ and $V =_{df} V' =_{df} V''$ are all mutually stochastically independent. This independence implies

$$
\begin{aligned}
F_{\min(A'' + B_A, V'' + B_V)}(t) &= F_{A'' + B_A}(t) + F_{V'' + B_V}(t) \\
&\quad - F_{A'' + B_A}(t)F_{V'' + B_V}(t) \\
&= F_{A + B_A}(t) + F_{V + B_V}(t) \\
&\quad - F_{A + B_A}(t)F_{V + B_V}(t).
\end{aligned}
\tag{22}
$$

Finally, we can substitute (22) into Eq. (21) to produce

$$
F_{\max(A', V') + B_{AV}}(t) \geqslant F_{A + B_A}(t)F_{V + B_V}(t)
$$

which is the required result, Eq. (18). $\square$

From Eq. (18) and some careful manipulation of logarithms it then follows that the measured capacity of a base time inclusive UCIP model is

$$
C_B^*(t) \leqslant 1.
$$

So the capacity of a UCIP model with base time is less than or equal to the capacity of a UCIP model without base time. Base time thus causes the capacity function to underestimate RSs facilitation and makes systems look less efficient than they really are. In Section 4.4 we estimate the size of this underestimation bias. In the meantime we briefly describe an experimental means of avoiding the bias altogether.

4.3. Toward an unbiased test

The preceding considerations have demonstrated that base time can alter the diagnosticity of the capacity coefficient; the test of UCIP models is weakened from an

equality to an inequality. We briefly present some ideas about the possibilities for recovering race-model tests that are unbiased by base time.

The new test is based on the earlier observation that, regardless of the dependence between A and V,

$$F_{A+B_A}(t) + F_{V+B_V}(t) = F_{\min(A',V')+B_{AV}}(t)$$
$$+ F_{\max(A'',V'')+B_{AV}}(t). \quad (19'')$$

In order to make practical use of this equality we will need to augment our standard RSs experiment with a condition in which subjects have to monitor *both* modalities in order to make a correct decision (Townsend & Fifić, 2004). This condition allows us to estimate the quantity $F_{\max(A,V)+B_{AV}}(t)$, since subjects must now process both channels and it is the *slower* of the two channels that determines the time taken for the decisional process to complete. In other words we propose the following mapping from experimental to theoretical variables:

Audio SS trial : $RT_A = A + B_A,$
Visual SS trial : $RT_V = V + B_V,$
Audiovisual redundant trial :
                         $RT_{AV}^{\mathrm{OR}} = \min(A', V') + B'_{AV},$
Audiovisual exhaustive trial :
                         $RT_{AV}^{\mathrm{AND}} = \max(A'', V'') + B''_{AV}.$
$$\quad (23)$$

In order to make this test work we will have to make the strong assumption that subjects employ the same channels A and V across both redundant signals (OR) and exhaustive processing (AND) tasks, and that the dependencies between A and V are also the same across these conditions. This assumption, which can be understood as an extension of standard *context invariance*, is most likely to hold when A and V are processed independently; it is then not unrealistic to assume that modalities processed independently in an OR task are again processed independently in an AND task. If we accept this step, then we can define a new capacity function:

$$C_B(t) = \frac{\log[P(RT_{AV}^{\mathrm{OR}} > t)]}{\log[P(RT_A > t) + P(RT_V > t) - P(RT_{AV}^{\mathrm{AND}} > t)]}. \quad (24)$$

This new capacity function can be interpreted in just the same manner as $C(t)$ (in terms of super, unlimited and limited capacity), but it provides predictions that are invariant under the inclusion/exclusion of base time from the processing system. A closely analogous $K_B(t)$ function can be defined that extends the $K(t)$ function currently used in tests of AND processing architecture (Townsend & Wenger, 2004). We do not elaborate on this possibility any further here. Instead we move on to consider how large an effect the base time is having on our present capacity coefficient, $C(t)$.

4.4. Analytic estimates of base time influence on capacity

4.4.1. Exponentially distributed base time

In order to quantitatively estimate the magnitude of the these base time effects, we must make some more specific assumptions about the processes A, V and B. For the sake of simplicity we will assume in this first analysis that all the processes in our system (cognitive and non-cognitive) are exponentially distributed. There are several advantages in assuming an exponential distribution:

(1) It is a common component of many process models going back to Hohle (1965), McGill (1963, Stochastic latency mechanisms) and Christie and Luce (1956).
(2) There is strong evidence that it closely approximates the distributions of inserted items in search tasks (Ashby, 1982), although statistical mimicking from other distributions can occur in lower power data (Ratcliff, 1988).
(3) Its simple functional form makes tractable the derivation of analytic formula that defines the properties of our model system. We therefore do not have to perform stochastic simulations which are often very imprecise in their predictions regarding low-probability events.

It is however, true, that an all-exponential model does not fit empirical data well; we will consider a somewhat more realistic base-time-inclusive model in Section 4.4.2.

Suppose then that $A, A' \sim \mathrm{Exp}(\alpha)$, $V, V' \sim \mathrm{Exp}(\upsilon)$ and $B \sim \mathrm{Exp}(\beta)$. The *survivor function*, $S(t)$ of the variable $A + B_A$ can then be written

$$S_{A+B_A}(t) = P(A + B_A \geqslant t)$$
$$= 1 - F_{A+B_A}(t)$$
$$= 1 - \int_{-\infty}^{t} f_A(t-s)F_B(s)\,ds$$
$$= \frac{\beta e^{-\alpha t} - \alpha e^{-\beta t}}{\beta - \alpha}, \quad (25)$$

where F_B is the common cdf of the base time components B_A, B_V and B_{AV}.

Similarly we obtain

$$S_{V+B_V}(t) = P(V + B_V \geqslant t)$$
$$= \frac{\beta e^{-\upsilon t} - \upsilon e^{-\beta t}}{\beta - \upsilon}.$$

It is not difficult to show that if $A' \sim \mathrm{Exp}(\alpha)$ and $V' \sim \mathrm{Exp}(\upsilon)$ then $\min(A', V') \sim \mathrm{Exp}(\alpha + \upsilon)$. From this it follows that for the RSs trials we have

$$S_{\min(A',V')+B_{AV}}(t) = P(\min(A', V') + B_{AV} \geqslant t)$$
$$= 1 - F_{\min(A',V')+B_{AV}}(t)$$
$$= 1 - \int_{-\infty}^{t} f_{\min(A',V')}(t-s)F_B(s)\,ds$$
$$= \frac{\beta e^{-(\alpha+\upsilon)t} - (\alpha+\upsilon)e^{-\beta t}}{\beta - (\alpha+\upsilon)}. \quad (26)$$

Using these expressions for the survivor functions we can now write

$$
\begin{aligned}
C(t) &= \frac{\ln S_{\min(A',V')+B_{AV}}(t)}{\ln S_{A+B_A}(t) + \ln S_{V+B_V}(t)} \\
&= \frac{\ln S_{\min(A',V')+B_{AV}}(t)}{\ln S_{A+B_A}(t) S_{V+B_V}(t)} \\
&= \frac{\ln\left[\dfrac{\beta e^{-(\alpha+\upsilon)t} - (\alpha+\upsilon)e^{-\beta t}}{\beta-(\alpha+\upsilon)}\right]}{\ln\left[\dfrac{(\beta e^{-\alpha t} - \alpha e^{-\beta t})(\beta e^{-\upsilon t} - \upsilon e^{-\beta t})}{(\beta-\alpha)(\beta-\upsilon)}\right]}
\end{aligned}
\tag{27}
$$

for the RSD with base time included in which all processes are exponentially distributed.

In Fig. 11 we can see Eq. (27) plotted with fixed channel process rates ($\alpha = \upsilon = 1$) but varying base time ($\beta \in [0,5]$) and for time a span that is large relative to the mean process time ($t \in [0,10]$).

When β is *large* relative to α and υ, i.e. when the mean and variance of the base time are *small* relative to the total reaction time the capacity index does not deviate far from the desired $C(t) = 1$, even for large t.

But when β is *small* relative to α and υ, i.e. when mean and variance of the base time are *large* relative to the total reaction time, the capacity index does indeed deviate from $C(t) = 1$. The deviation from $C(t) = 1$ down to $C(t) = \frac{1}{2}$ is a large deviation; bear in mind that a capacity coefficient $C(t) = \frac{1}{2}$ is what we would normally expect from minimum-time serial processing, where if $H(t) = H_V(t) = H_A(t) = H_{AV}(t)$, then it follows trivially that $C(t) = \frac{1}{2}$. This is a special case of the well-known *fixed capacity* parallel model. Under these conditions, it is also the model which forms the edge of the Grice bound, $F_{AV}(t) = \max[FA(t), FV(t)]$ (see Colonius, 1990; Grice, Canham, & Boroughs, 1984; Townsend & Nozawa, 1997).
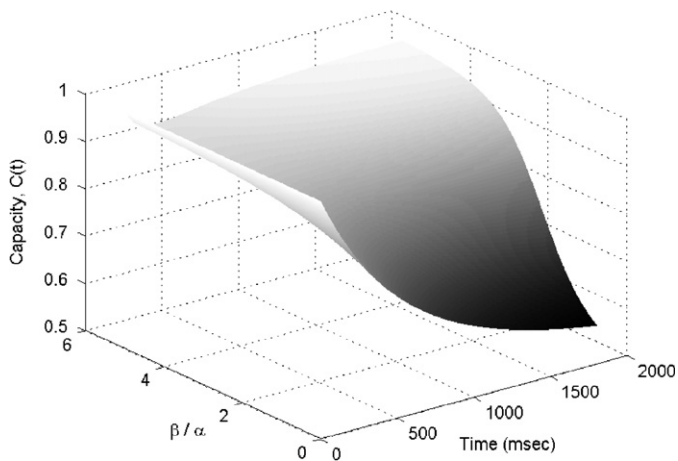

Fig. 11. Plot of capacity $C(t)$ measured from a UCIP model with base time. A, V and $B_{AV} =_{df} B_A =_{df} B_V$ are all exponentially distributed. The common mean and standard deviation of the base time variables are varied by varying the exponential parameter β. The process variables had fixed exponential parameters $\alpha = \upsilon = \frac{1}{450}$ ms.

Although Fig. 11 does show the *possibility* of large deviations of $C(t)$ from what we would expect in the absence of base time, the range of parameters that produce these large deviations appears unrealistic. In our discussion in Section 5 we marshal some evidence (Evarts, 1966; Kohfeld et al., 1981; Meijers & Eijkman, 1974; Ratcliff, Thapar, & McKoon, 2001) indicating that the variance of the base time is typically much smaller than the variance of the channel process variables A and V. In terms of our model parameters this means that we are most interested in the parameter regime in which β is greater than $\alpha(= \upsilon)$. The deviations shown in Fig. 11 for the regime $\beta/\alpha > 1$ are not large; and when $\beta/\alpha > 2$ (i.e when the standard deviation of the base time is less than twice the standard deviation of the process time) then the deviations are negligible.

Experimenters should remain wary of paradigms in which they can reasonably believe that a large proportion of the total reaction time variance is due to variance in the base time. However, as we argue in more detail in Section 5, this is unlikely in the majority of RSDs, and the impact of base time on the diagnosticity of the capacity coefficient (as well as the race model inequality) is therefore likely to be small.

4.4.2. Normally distributed base time

The results from the previous section can be criticized on the grounds that the mean and variance of the exponential distribution are entirely confounded, since they are both functions of the same parameter, β. The analysis presented above cannot dissociate effects due to large-*variance* base time from effects due to large-*mean* base time.

We therefore turn to consider the case in which channel processes, A and V, follow an exponential distribution but the base time, B, is distributed normally (see Heathcote, Popiel, & Mewhort, 1991; Luce, 1986; Ratcliff, 1976). Introducing the Normal distribution will complicate the calculations somewhat, but will grant us the freedom to independently vary the mean and variance of B.

So we now have $A, A' \sim \text{Exp}(\alpha)$, $V, V' \sim \text{Exp}(\upsilon)$ and $B_A, B_V, B_{AV} \sim \text{N}(\mu, \sigma)$. Recall that if $G \sim \text{Exp}(\lambda) + \text{N}(\mu, \sigma)$ then we say that G follows an ex-Gaussian distribution: $\text{ExG}(\lambda, \mu, \sigma)$. The random variables $A + B_A$ and $V + B_V$ will therefore follow ex-Gaussian distributions. In Appendix C we provide a derivation of the cumulative distribution function of the ex-Gaussian:

$$
\begin{aligned}
F_{\text{ExG}}(t) = {}& -e^{-\lambda(t-\mu)+\sigma^2\lambda^2} \Phi\left(\frac{t-\mu-\sigma^2\lambda}{\sigma}\right) \\
& + \Phi\left(\frac{t-\mu}{\sigma}\right).
\end{aligned}
\tag{28}
$$

So now we have

$$
\begin{aligned}
S_{A+B_A}(t) &= P(A + B_A \geqslant t) \\
&= 1 - F_{A+B_A}(t) \\
&= 1 + e^{-\alpha(t-\mu)+\sigma^2\alpha^2} \Phi\left(\frac{t-\mu-\sigma^2\alpha}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right),
\end{aligned}
$$

$$\begin{aligned} S_{V+B_V}(t) &= P(V + B_V \geqslant t) \\ &= 1 - F_{V+B_V}(t) \\ &= 1 + e^{-v(t-\mu)+\sigma^2 v^2} \Phi\left(\frac{t-\mu-\sigma^2 v}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right), \end{aligned}$$

$$\begin{aligned} S_{\min(A',V')+B_{AV}}(t) &= P(\min(A',V') + B_{AV} \geqslant t) \\ &= 1 - F_{\min(A',V')+B_{AV}}(t) \\ &= 1 + e^{-(\alpha+v)(t-\mu)+\sigma^2(\alpha+v)^2} \Phi \\ &\quad \times \left(\frac{t-\mu-\sigma^2(\alpha+v)}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right). \end{aligned}$$

$$(29)$$

Substituting these into our capacity measure, Eq. (15), we obtain

$$C(t) = \frac{\ln\left[1 + e^{-(\alpha+v)(t-\mu)+\sigma^2(\alpha+v)^2} \Phi\left(\frac{t-\mu-\sigma^2(\alpha+v)}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right)\right]}{\ln\left[\left(1 + e^{-\alpha(t-\mu)+\sigma^2\alpha^2} \Phi\left(\frac{t-\mu-\sigma^2\alpha}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right)\right)\left(1 + e^{-v(t-\mu)+\sigma^2 v^2} \Phi\left(\frac{t-\mu-\sigma^2 v}{\sigma}\right) + \Phi\left(\frac{t-\mu}{\sigma}\right)\right)\right]}. \qquad (30)$$

With these two quantitative capacity measures in hand, we can begin to investigate the effects of base time on capacity, at least in this restricted domain.

In Fig. 12 we can see Eq. (30) plotted with

- the exponentially distributed channel processes with rates fixed at $\alpha = v = \frac{1}{500}\,\mathrm{ms}^{-1}$.
- the normally distributed base time process having a relatively small mean $\mu = 100\,\mathrm{ms}$ and standard deviation $\sigma \in [0, 60]\,\mathrm{ms}$.
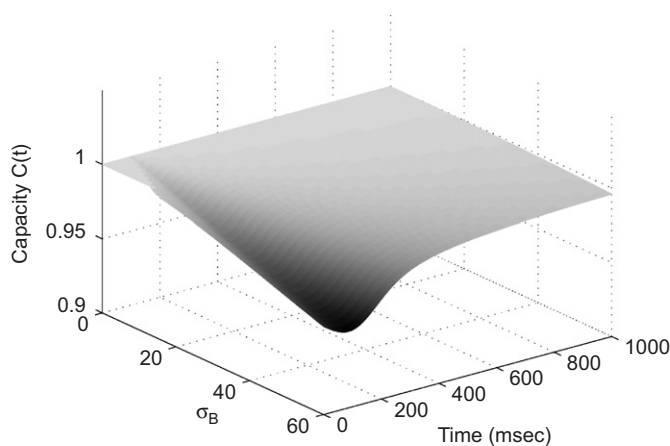- over the time window $t \in [0, 1000]$.

Fig. 12. Plot of capacity $C(t)$ measured from a UCIP model with base time. A and V are exponentially distributed; the base time variables are normally distributed. The common mean of the base time variables is $\mu = 100$; the standard deviation σ_B of the base time varies between 0 and 60 ms. The process variables had fixed exponential parameters $\alpha = v = \frac{1}{500}\,\mathrm{ms}^{-1}$.

In our investigations, varying the common mean of the base time variables had negligible impact on the shape of $C(t)$. Changes in $C(t)$ were driven by the base time standard deviation, σ, and so this is the parameter we vary in Fig. 12. Of course this pattern is in accord with intuition: a large constant base time (i.e. a base time with zero variance) is far less of a problem for the experimentalist than a base time which has a small mean but which is producing most of the observed variance in reaction times.[14] As pointed out in Section 3.2, the mean of the base time determines the magnitude of the horizontal shift in our reaction time cdf while the variance of the base time determines the magnitude of the smoothing effect.

We see again that—just as with the race model inequality—base time of proportionally large variance does lead the capacity coefficient to *underestimate* facilitatory effects in data. But this analysis using normally distributed base time—even more so than the analysis in the previous section using exponentially distributed base time—suggests that the effects of base time on the capacity coefficient are negligible in the parameter range of almost any RSD. This finding may offer comfort to the experimentalist employing $C(t)$ to measure capacity.

5. Discussion

5.1. Varying views of base time

We must refrain from hastily labeling any brain process as "non-cognitive". Visual-system information is processed even at the retinal level, and top–down influences on basic sensory processes are well documented (e.g. Crist, Li, & Gilbert, 2001). But it seems plausible that by varying the salience of experimental stimuli we can selectively facilitate some of the processes involved in task performance while leaving other processes—such as the motor response time—unchanged. It is in this manner that we operationally define which are the cognitive/processing components

---

[14] Another thing to note in the figure is that for base time with non-zero variance we observe $C(0) < 1$. This does not seem to make sense, because at $t = 0$ all cdf's should be zero since they are cdf's of strictly positive random variables; the numerator and denominator of $C(t)$ should thus be equal and so we should have $C(0) = 1$ regardless of any other factor. The reason for this apparently anomalous behavior is that since base time follows a Normal distribution it *can* take on negative values; so our reaction times are no longer strictly positive random variables and—in particular with high-variance base time—we can observe $C(0) < 1$. This is not a matter of great significance; the effects are negligible for realistic values of the mean and standard deviation of base time.
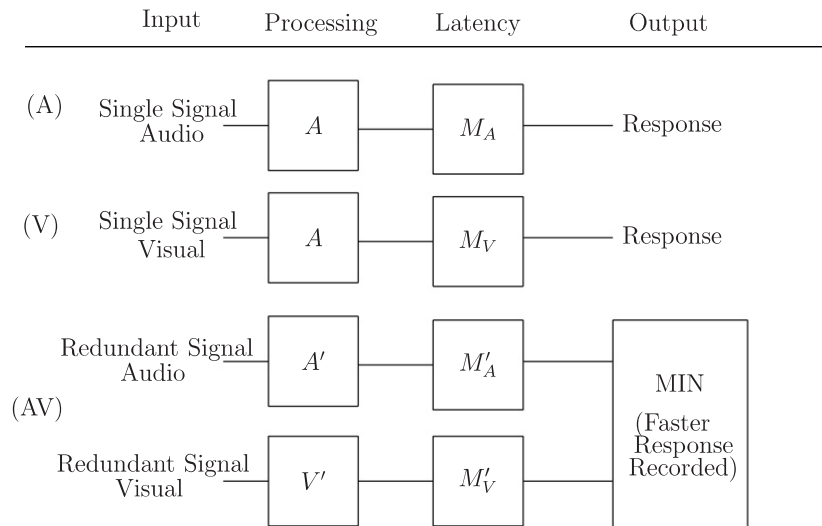
Fig. 13. Here we attempt to reconstruct a model that produces the result that $RT_{AV} = \min(A' + B_A, V' + B_V)$. In this conception of base time (Maris & Maris, 2003) both channels can initiate a motor process and the "minimum" operation is implicitly generated on RS trials because the response time measuring apparatus will only record the response from the channel which first completes both detection *and* motor processes.

and which are the non-cognitive/base components of a reaction time.

Base time is then defined as that component of our measured reaction times that is

- independently and identically distributed across trials,
- independently and identically distributed across experimental conditions.

By definition, then, our base time is equal in distribution across conditions ($B_A =_{df} B_V =_{df} B_{AV}$) and the base time variables are statistically independent of the process variables A and V.

Before we move on to discuss the likely theoretical and empirical approaches that will help us to understand the significance of base time for the RSs task and for reaction time paradigms in general, we must address some possible criticisms of the model of base time that we present in this paper. There is, on the one hand, some doubt about the *existence* of a base time component that is independent of processing time, and on the other hand, dispute about the architectural *locus* of this component.

By varying stimulus intensity in a simple visual displacement–detection task Dzhafarov (1992) found evidence for the thesis that one cannot consistently decompose a reaction time into two components, one independent of and the other dependent upon stimulus intensity. He proposed an alternative model in which process time and base time are *both* covariates of a single "readiness factor" that varies stochastically between trials.[15] This kind of result can also be explained by a

processing network comprising discrete time random variables that are perfectly correlated.

When variables in a process model become highly correlated with one another it becomes awkward to continue modeling them as separate entities. Nevertheless, motor response and target detection are unlikely to be integrated processes when the experimental task imposes decisional demands that are more complex than those in the visual displacement–detection task that Dzhafarov (1992) reports.[16]

Maris and Maris (2003) accept the standard account of the existence of independent base time components but they suggest that there is an acceptable alternative account of the location of base time in the cognitive architecture. In their alternative conception, the two channel process times A' and V' each have associated base times M'_A and M'_V (see Fig. 13) but there is no 'post-decisional' base time.[17] This implies that the audio and decisional processes have independent access to the initiation of the motor response, and that the winner of the "race" in the RS condition is just the channel that first completes its processing *and* its motor response, i.e. the channel that produces the experimentally observed button-push. The authors point out that "With

---

[15]However, Ratcliff (personal communication) points out that across numerous papers and large data sets he has found no correlation between boundary separation and base time. To the extent that the boundary separation parameter in the diffusion model framework is equivalent to

(*footnote continued*)
the "readiness factor" mentioned above, these results are incompatible with those of Dzhafarov (1992).

[16]Consider, for example, a task in which subjects are instructed to add two integers—one presented in the audio and the other presented in the visual modality—and to respond when the sum of the integers was equal to, say, 16. It is difficult to imagine this task being performed in the absence of functional stages.

[17]As far as we see this arrangement is the only way to interpret their suggestion that reaction times in the RS condition are described by the variable $\min(A' + M'_A, V' + M'_V)$ as opposed to the more standard $\min(A', V') + M_{AV}$. If the base time variable occurs only inside the minimum operator then there cannot be any base time component in effect after the "race" is completed.
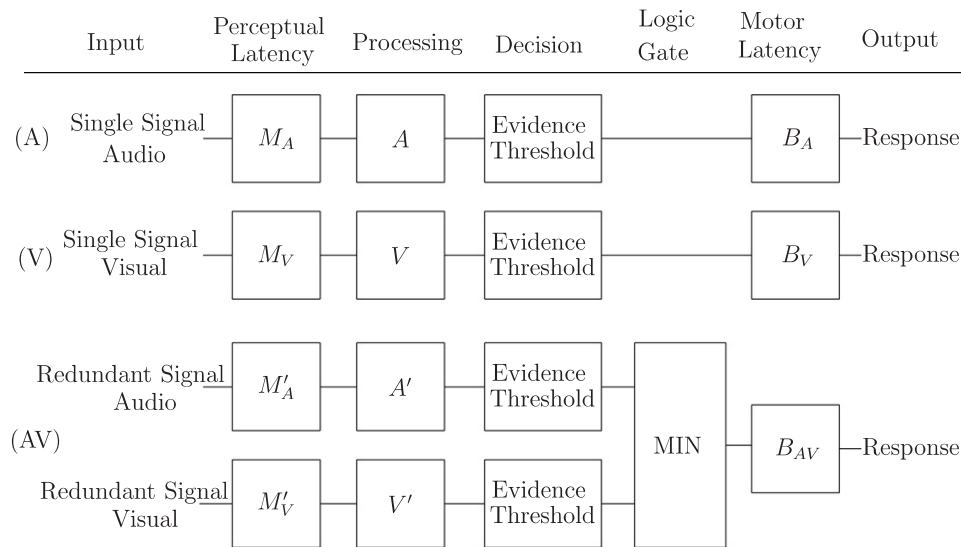
Fig. 14. Schematic of a general redundant signals model incorporating both 'local' and 'global' base times.

this assumption, the difference between response time and decision time vanishes, in the sense that we cannot disentangle the two components of the response time." (Maris & Maris, 2003). However, a great many studies (Ilan & Miller, 1998; Pashler & Johnston, 1998, Attentional limitations in dual-task performance; Sommer, Leuthold, & Schubert, 2001) support the concept of two discrete stages: a bottleneck response selection phase followed by a motor execution phase.[18]

We agree that there may be non-cognitive elements *within* each process channel. Consider the more general model shown in Fig. 14. In this model two types of base time have been included: a "global" base time and a "local" time. In the SS condition each channel (audio, A, and visual, V) has its own local base time (M_A and M_V) which models the most basic perceptual latencies and there is also a 'global' base time (B_A and B_V) which corresponds roughly to the motor component of the reaction time. In the RSs condition we have an analogous set-up, but now a common motor component B_{AV} is located after the decisional bottleneck. It is our contention that we can

safely neglect the variables M_A, M'_A and M_V, M'_V because they occur in series with A, A' and V, V', respectively and will therefore always be confounded with those two variables. In choosing to neglect M_A and M_V, then, we are defining each channel process variable to include both processing time and "local" base time. This is equivalent to writing $A \equiv M_A + A$, $A' \equiv M'_A + A'$, $V \equiv M_V + V$, $V' \equiv M'_V + V'$, and thereby recovering the standard model shown in Fig. 2 and employed throughout this paper.

The uncertainty about the place and the properties of base time reflects a more general difficulty in linking behaviorist-style experimental data with the insights from more recent approaches based on EEG (Ilan & Miller, 1998; Sommer et al., 2001) and fMRI (Schubert & Szameitat, 2003). The behavioral and neurological approaches can undoubtedly be implemented synergistically, but mapping results from one field onto those of the other without understanding the differences in terminology and methodology will simply result in confusion. As Sternberg (2001) points out: "A stage is a function carried out during an epoch in time; it is not necessarily associated with a distinct neuroanatomical processor. Too often, an inferred flowchart, which describes the temporal arrangement of a set of operations (and is inherently ordered, because of the nature of time), has been confused with a circuit diagram, which indicates how processors are connected."

Given the instrumental definition of base time that we suggest above, it may well emerge that the activation time-course of neurons in *motor cortex* is, perplexingly, uncorrelated with base time and yet strongly correlated with the process times on the detection channels. This could be the case if, for example, all of the process-independent variability in motor time is a consequence of neural noise in the basal ganglia or the spinal column, two areas inaccessible to EEG recording.

However, single-cell recording data from monkeys (Evarts, 1966) indicates (a) that "motor time" and "pre-motor time"

---

[18]Naturally this is an empirical question which we cannot settle here. But it should be possible to decide between these two pictures of base time on empirical grounds. For instance, consider systems factorial technology, based on selective influence of experimental factors on distinct sub-processes (e.g. Schweickert, 1978; Sternberg, 1969; Townsend & Nozawa, 1995). If the model of Fig. 2 is correct, it should be straightforward to find an additive factor associated with the motor time. This should be impossible with the model of Fig. 13. In principle, a RSs experiment in which subjects respond to the audio signal and to the visual signal with separate response-buttons might even help decide the issue. If the model that employs a shared post-decisional base time is accurate, then we would expect that experimental subjects would very infrequently respond by pressing both the audio response and the visual response buttons. If the model that employs independent motor responses is correct then we would expect more of these double-presses. Even as a thought experiment, it seems unlikely these predictions would be verified. For an example of experimentation involving the manipulation of response modalities see Hughes (1994).

are uncorrelated in a hand-movement task and (b) that the standard deviation of motor time was 8 ms. Meijers and Eijkman (1974) pointed out that, based on the same single-cell recording data, "the ratio between the variance of premotor-time and the variance of the motor-time varied from 5–10". These estimates of base time are in accord with the data of Kohfeld et al. (1981) who report a mean of 136 ms and a standard deviation of 9.1 ms for the base time.

The analyses we present in this paper indicate that the crucial factor producing bias in our model-selection tools is the proportion of reaction time variance contributed by variance in the base time and the quantitative estimates we cite in the paragraph above are therefore reassuring. However, there are also larger estimates of base time variance in the literature. Based on parameter fits to a diffusion model incorporating a uniformly distributed base time, Ratcliff and Tuerlinckx (2002) and Ratcliff and Smith (2004) report that the base time variable is located within an absolute range of 100–200 ms around its mean. These absolute range estimates corresponds to standard deviations of approximately 30–60 ms. If the standard deviation of the base time does indeed fall at the higher end of this larger range then the smoothing effects can in fact be quite substantial (see, e.g. Fig. 5) and the diagnosticity of our model selection tools can be seriously compromised.

One explanation for the different values of base time reported in the literature is that base time is not an identical quantity across experimental conditions and tasks. Base time may well have a larger mean and variance in more complicated tasks, and this possibility underlines the uncertainty in our very definition of this theoretical construct. However, in more complex tasks the increase in base time variance is likely to be matched by an increase in the variance of the process times. In this case the proportion of variance contributed by base time, and the impact of base time on our model selection tools, would remain small. We eagerly anticipate the marriage of theoretical approaches such as Fourier deconvolution (Goldstone, 2000, but note Sheu & Ratcliff, 1995) and formal non-parametric tests (Cortese & Dzhafarov, 1996; Van Zandt, 2002, Analysis of response time distributions) with novel neuroimaging techniques toward an understanding and resolution of the base time problem.

6. Conclusions

We have demonstrated the relationship between the race model test function $R^*(t)$ and the capacity coefficient function $C(t)$. We provide distribution-free analytic results indicating that both of these model-selection tools will be biased when base time is present.

Typically, the presence of base time produces a bias toward underestimation of RSs facilitation in both R^* and $C(t)$. The magnitude of this model-selection bias increases as we increase the proportion of total reaction time variance that is contributed by variance in base time. Additionally, we investigated the statistical power of model

selection tools based on the race model inequality using Monte Carlo simulations and showed that the effects of base time on power are usually small. We also sketched out a means of building a statistical test function which, relying on data from both AND tasks and OR tasks, would be unbiased in the presence of base time.

Finally, we provided quantitative estimates of the effects of base time for a range of realistic reaction time distributions. In general, the timing and magnitude of the base times bias effects depend non-trivially on the specific distributions governing the channel and base time processes. Under relatively realistic assumptions and within the regime of standard parameters as described in the text, it appears that bias effects will be minor. We note, however, that uncertainty remains concerning the magnitude and even the definition of base time, and suggest that experimentalists continue to be alert to the possibility of data contamination from this source.

Appendix A. A note on notation

Random variable notation must be handled carefully when describing mathematical models of RSs effects, because the models commonly involve variables being realized marginally in one context and jointly in another. It is important to be clear about which kind of variable realization is taking place. For convenience we assume the standard convention under which multiple instances of the same random variable in a single equation denote the same realization of that random variable because it allows us to write down intuitive formulations such as $X + X = 2X$.

When modeling data from a RSs experiment—an experiment in which reaction time data for the A, V and AV are gathered on separate trials or even on separate days of experimentation—we must therefore be sure to use different random variables to describe the processes ongoing in each condition. When context invariance is assumed many of the variables across conditions are identically distributed and we can be tempted to denote these separate variables with a common symbol, but this can lead us into subtle mistakes. This occurs, for example, in the Appendices of Ulrich and Giray (1986), where the authors attempt to theoretically link their SS and RSs data by means of the following manipulation:

$$RT_{AV} = \min(A, V) + B$$
$$= \min(A + B, V + B)$$
$$= \min(RT_A, RT_V).$$

The problem with this reasoning is that RT_A and RT_V are variables modeling data gathered on entirely separate occasions and yet they are being substituted for the terms $A + B$ and $V + B$ which share a common factor B that always realizes to an identical value in both expressions. One way to avoid making this kind of error is to use the more careful (and unfortunately less elegant) notation in which distinct (but identically distributed) base time variables B_A, B_V and B_{AV} are deployed for each condition.[19] This is a small point, but one with the potential to cause confusion. In Appendix B we provide a corrected proof of Ulrich and Giray's Theorem 1.

Appendix B. Revised proof

The following is a revision of Theorem 1 from Ulrich and Giray (1986). For an elegant version of the RHS result see Luce (1986, p. 131); this book contains a comprehensive and insightful commentary on the base time issue, and offers the reader more perspective than we can provide here.

Theorem 7. *If X and Y are (possibly dependent) non-negative channel process variables and $B_X =_{df} B_Y =_{df} B_{XY}$ are non-negative base time variables that are stochastically independent of X and Y, then*

$$\max[F_{X+B_X}(t), F_{Y+B_Y}(t)] \leqslant F_{\min(X',Y')+B_{XY}}(t)$$
$$\leqslant F_{X+B_X}(t) + F_{Y+B_Y}(t).$$

Proof. We prove the leftmost inequality first. We begin by adopting the notation $f_B(t)$ for the common density function of the variables B_X, B_Y and B_{XY} and simply writing

$$F_{\min(X',Y')+B_{XY}}(t) - \max[F_{X+B_X}(t), F_{Y+B_Y}(t)]$$
$$= \int_{-\infty}^{t} F_{\min(X',Y')}(s) f_B(t-s)\, ds$$
$$- \max \left\{ \int_{-\infty}^{t} F_X(s) f_B(t-s)\, ds, \right.$$
$$\left. \times \int_{-\infty}^{t} F_Y(s) f_B(t-s)\, ds \right\}$$

at which point we can choose, without loss of generality, the value of the max function on the RHS. We then have

$$F_{\min(X',Y')+B_{XY}}(t) - \max[F_{X+B_X}(t), F_{Y+B_Y}(t)]$$
$$= \int_{-\infty}^{t} (F_{\min(X',Y')}(s) - F_X(s)) f_B(t-s)\, ds$$
$$= \int_{-\infty}^{t} [\text{Positive Fn}][\text{Positive Fn}]\, ds$$
$$\geqslant 0 \tag{B.1}$$

demonstrating the required result, which is essentially a base time generalization of the Grice Inequality (Grice et al., 1984).

We now proceed to prove the rightmost inequality, $F_{\min(X',Y')+B_{XY}}(t) \leqslant F_{X+B_X}(t) + F_{Y+B_Y}(t)$. We begin by considering the difference of the two quantities:

$$F_{\min(X',Y')+B_{XY}}(t) - F_{X+B_X}(t) - F_{Y+B_Y}(t)$$
$$= \int_{-\infty}^{t} F_{\min(X',Y')}(s) f_B(t-s)\, ds$$
$$- \int_{-\infty}^{t} F_X(s) f_B(t-s)\, ds$$
$$- \int_{-\infty}^{t} F_Y(s) f_B(t-s)\, ds$$
$$= \int_{-\infty}^{t} [F_{\min(X',Y')}(s) - F_X(s) - F_Y(s)] f_B(t-s)\, ds. \tag{B.2}$$

Now we recall that $F_{\min(X',Y')}(s) - F_X(s) - F_Y(s) \leqslant 0$ from the standard version of the race model inequality, Eq. (3), so that we can rewrite the above as

$$F_{\min(X',Y')+B_{XY}}(t) - F_{X+B_X}(t) - F_{Y+B_Y}(t)$$
$$= \int_{-\infty}^{t} [\text{Negative Fn}][\text{Positive Fn}]\, ds$$
$$\leqslant 0 \tag{B.3}$$

which provides the desired result

$$F_{\min(X',Y')+B_{XY}}(t) \leqslant F_{X+B_X}(t) + F_{Y+B_Y}(t). \tag{B.4}$$

This completes the proof of the rightmost inequality, which is a base time generalization of the Miller Race Model Inequality. $\square$

Appendix C. The ex-Gaussian distribution

Recall that the probability density function of the ex-Gaussian distribution is given by

$$f_{\text{ExG}}(t) = \frac{1}{\tau} e^{(-t+\mu)/\tau + \sigma^2/2\tau^2} \Phi\left(\frac{t - \mu - \frac{\sigma^2}{\tau}}{\sigma}\right),$$

where ϕ is the cumulative distribution function of a Gaussian distribution with $\mu = 0, \sigma = 1$.

Now $F_{\text{ExG}}(t) = \int_{-\infty}^{t} f_{\text{ExG}}(s)\, ds$ so we need to integrate the pdf above to obtain the cdf. First we recall the notation

$$\Phi(x) = \int_{-\infty}^{x} e^{-s^2/2}\, ds$$

which we will employ as we *integrate by parts*, i.e. we use the fact that

$$\int_{-\infty}^{t} f(s) g'(s)\, ds = [f(s)g(s)]_{-\infty}^{t} - \int_{-\infty}^{t} f'(s)g(s)\, ds$$

---

[19]Colonius and Vorberg (1994) make a similar slip when adding a common motor component M to a variety of separate decisional variables (pp. 44–45).

to transform our integral into a more tractable form.

$$F_{\text{ExG}}(t) = \int_{-\infty}^{t} f(s)\,ds$$

$$= \int_{-\infty}^{t} \frac{1}{\tau} e^{(-s+\mu)/\tau+\sigma^2/2\tau^2} \Phi\left(\frac{s-\mu-\frac{\sigma^2}{\tau}}{\sigma}\right) ds$$

$$= \left[-e^{(-s+\mu)/\tau+\sigma^2/\tau^2} \Phi\left(\frac{s-\mu-\frac{\sigma^2}{\tau}}{\sigma}\right)\right]_{-\infty}^{t}$$

$$+ \int_{-\infty}^{x} e^{((-s+\mu)/\tau+\sigma^2/2\tau^2)} e^{-(1/2)(t-\mu-\sigma^2/\tau/\sigma)^2}\,ds$$

$$= -e^{(-t+\mu)/\tau+\sigma^2/\tau^2} \Phi\left(\frac{t-\mu-\frac{\sigma^2}{\tau}}{\sigma}\right)$$

$$+ \int_{-\infty}^{t} e^{-(1/2)((s-\mu)/\sigma)^2}\,ds$$

$$= -e^{(-t+\mu)/\tau+\sigma^2/\tau^2} \Phi\left(\frac{t-\mu-\frac{\sigma^2}{\tau}}{\sigma}\right)$$

$$+ \Phi\left(\frac{t-\mu}{\sigma}\right).$$

It is not difficult to see that $F(t)$ is an increasing function with the properties $\lim_{t\to-\infty} F(t) = 0$ and $\lim_{t\to\infty} F(t) = 1$; it is therefore also a cumulative distribution function as required.

Appendix D. Notations and abbreviations

Table D.1 shows a complete list notations and abbreviations.

Table D.1
Notations and abbreviations used in the text

Symbol	Meaning	Note
SS	Single signal	Experimental condition
RS	Redundant signals	Experimental condition
UCIP	Unlimited capacity, independent, parallel	Type of race model
cdf	Cumulative distribution function	
A	Process time in SS audio condition	Theoretical random variable
V	Process time in SS visual condition	Theoretical random variable
AV	Process time in RS audio-visual condition	Theoretical random variable
B_A	Base time in SS audio condition	Theoretical random variable
B_V	Base time in SS visual condition	Theoretical random variable

Table D.1 (*continued*)

Symbol	Meaning	Note
B_{AV}	Base time in RS audio-visual condition	Theoretical random variable
RT_A	Reaction time in SS audio condition	Empirical random variable
RT_V	Reaction time in SS visual condition	Empirical random variable
RT_{AV}	Reaction time in RS audio-visual condition	Empirical random variable
$=_{df}$	'is equal in distribution'	Relation between random variables
X'	Random variable with property $X'=_{df} X$	
$F_X(t)$	Marginal cdf of X	$= P(X \leqslant t)$
$S_X(t)$	Marginal survivor function of X	$= P(X > t)$
$f_B(t)$	Marginal probability density of B_A, B_V, B_{AV}	
$C(t)$	Capacity function	See Townsend and Nozawa (1995)
$C^*(t)$	Capacity function under UCIP assumptions	

References

Ashby, F. G. (1982). Testing the assumptions of exponential additive reaction time models. *Memory & Cognition*, *10*, 125–134.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*(2), 154–179.

Christie, L. S., & Luce, R. D. (1956). Decision structure and time relations in simple choice behavior. *Journal of Mathematical Biophysics*, *18*, 89–112.

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

Colonius, H., & Diederich, A. (2006). The race model inequality: Interpreting a geometric measure of the amount of violation. *Psychological Review*, *113*(1), 148–154.

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

Cortese, J. M., & Dzhafarov, E. N. (1996). Empirical recovery of response time decomposition rules II. Discriminability of serial and parallel architectures. *Journal of Mathematical Psychology*, *40*, 203–218.

Crist, R. E., Li, W., & Gilbert, C. D. (2001). Learning to see: Experience and attention in primary visual cortex. *Nature Neuroscience*, *4*, 519–525.

Donders, F. C. (1869). Over de snelheid van psychische processen. onderzoekingen gedaan in het physiologisch laboratorium der utrechtsche hoogeschool, 1868–1869, Tweede Reeks, II, 92–120, Translated by Koster, W. G. In W. G. Koster (Ed.), Attention and performance II. *Acta Psychologica* 30, 412–431.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Nature Neuroscience*, *36*, 235–268.

Evarts, E. V. (1966). Pyramidal tract activity associated with a conditioned hand movement in the monkey. *Journal of Neurophysiology*, *29*, 1011–1027.

Goldstone, R. L. (2000). Unitization during category learning. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 86–112.

Grice, G. R., Canham, L., & Boroughs, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, 35, 451–463.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of response time distributions: An example using the stroop task. *Psychological Bulletin*, 109(2), 340–347.

Hohle, R. H. (1965). Inferred components of reaction time as a function of foreperiod duration. *Journal of Experimental Psychology*, 69, 382–386.

Hughes, H. C. (1994). Visual–auditory interactions in sensorimotor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, 20(1), 131–153.

Ilan, A. B., & Miller, J. (1998). On the temporal relations between memory scanning and response preparation. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5), 1501–1520.

Ingvalson, E. M., & Wenger, M. J. (2005). A strong test of the dual-mode hypothesis. *Perception & Psychophysics*, 67(1), 14–35.

Kecs, W. (1982). *The convolution product and some applications.* Dordrecht: Reidel.

Kohfeld, D. L., Santee, J. L., & Wallace, N. D. (1981). Loudness and reaction time: I. *Perception & Psychophysics*, 29, 535–549.

Luce, R. D. (1986). *Response times.* Oxford: Oxford University Press.

Maris, G., & Maris, E. (2003). Testing the race model inequality: A nonparametric approach. *Journal of Mathematical Psychology*, 47, 507–514.

McGill, W. J. (1963). In R. D. Luce, R. R. Busch, & R. R. E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 309–360). New York: Wiley.

Meijers, L. M., & Eijkman, E. (1974). The motor system in simple reaction time experiments. *Acta Psychologica*, 38, 367–377.

Miller, J. (1978). Multidimensionsal same–different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 4(3), 411–422.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247–279.

Miller, J. (2004). Exaggerated redundancy gain in the split brain: A hemispheric coactivation account. *Cognitive Psychology*, 49, 118–154.

Mordkoff, J. T., & Miller, J. (1993). Redundancy gains and coactivation with two different targets: The problem of target preferences and the effects of display frequency. *Perception & Psychophysics*, 52, 527–535.

Pashler, H., & Johnston, J. (1998). In H. Pashler (Ed.), *Attention* (pp. 155–189). London: Erlbaum.

Patching, G. R., & Quinlan, P. T. (2004). Cross-modal integration of simple auditory and visual events. *Perception & Psychophysics*, 66(1), 131–140.

Raab, D. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, 24, 574–590.

Ratcliff, R. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83(3), 190–214.

Ratcliff, R. (1988). A note on the mimicking of additive reaction time models. *Journal of Mathematical Psychology*, 32, 192–280.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2), 333–367.

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16, 323–341.

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

Schubert, T., & Szameitat, A. (2003). Functional neuroanatomy of interference in overlapping dual tasks: An fmri study. *Brain Research Cognitive Brain Research*, 17(3), 733–746.

Schwarz, W., & Ischebeck, A. (1994). Coactivation and statistical facilitation in the detection of lines. *Perception*, 23, 157–168.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a stroop task. *Journal of Mathematical Psychology*, 18, 105–139.

Sheu, D.-F., & Ratcliff, R. (1995). The application of Fourier deconvolution to reaction time data: A cautionary note. *Psychological Bulletin*, 118, 234–251.

Smith, P. L. (1990). Obtaining meaningful results from Fourier deconvolution of reaction time data. *Psychological Bulletin*, 108(3), 533–550.

Snodgrass, J. G. (1969). Foreperiod effects in simple reaction time: Anticipation or expectancy? *Journal of Experimental Psychology, Monograph*, 79, 1–19.

Sommer, W., Leuthold, H., & Schubert, T. (2001). Multiple bottlenecks in information processing? An electrophysiological examination. *Psychonomic Bulletin & Review*, 8(1), 81–88.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57(4), 421–457.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica*, 106, 147–246.

Townsend, J., & Fifić, M. (2004). Parallel versus serial processing and individual differences in high-speed search in human memory. *Perception & Psychophysics*, 66(6), 953–962.

Townsend, J., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003–1035.

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

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, 321–360.

Townsend, J. T., & Nozawa, G. (1997). Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review*, 104(3), 595–602.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, 33(3), 309–327.

Ulrich, R., & Giray, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, 39(4), 248–254.

Van Zandt, T. (2002). Stevens' handbook of experimental psychology. In J. T. Wixted (Ed.), *Methodology in experimental psychology* (3rd Ed., Vol. 4) (pp. 461–516). New York: Wiley.

Wenger, M. J., & Townsend, J. T. (2001). In M. J. Wenger, & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 229–284). London: Erlbaum.

Westendorf, D. H., & Blake, R. (1988). Binocular reaction times to conrast increments. *Vision Research*, 28(2), 355–359.