JOURNAL OF MATHEMATICAL PSYCHOLOGY 36, 235–268 (1992)

The Structure of Simple Reaction Time to Step-Function Signals

EHTIBAR N. DZHAFAROV

University of Illinois at Urbana-Champaign and The Beckman Institute for Advanced Science and Technology

Simple reaction time to a step-function signal of amplitude A can be additively decomposed into a signal-dependent component, $\mathbf{T}(A)$ (converging to zero as A increases), and a signal-independent component, $\mathbf{R}$. The two components, however, are not mutually independent, they both are increasing deterministic functions of a single random variable; $\mathbf{T}(A) = T(A, \mathbf{C})$ and $\mathbf{R} = R(\mathbf{C})$. $\mathbf{C}$ is interpreted as a "criterion" or "inhibition factor" controlling simultaneously "readiness to detect" and "readiness to respond." The model is not based on a priori distributional assumptions. RT percentiles of a given rank P computed across RT distributions for different values of A, have a deterministic structure, $\mathrm{RT}_P(A) = T(A, C_P) + R(C_P)$, where C_P is the Pth percentile of $\mathbf{C}$. Subtracting $\mathrm{RT}_P(A^*)$ for a very large A^* from $\mathrm{RT}_P(A)$, one gets an estimate of $T(A, C_P)$, the Pth percentile of $\mathbf{T}(A)$. Applying this to different values of P, one can a posteriori reconstruct the joint distribution of the two RT components. The asymptotic consequences of this model (for sufficiently large A) are corroborated by data on the RT to instantaneous displacements of a visual target.

1. INTRODUCTION

1.1. *Brief Overview*

Simple reaction time (RT) to a step-function signal tends to decrease as a function of signal amplitude. Assumptions underlying this statement are that (a) step changes in a stimulus parameter occur from a fixed base level; (b) response is required as soon as a change is detected (simple RT paradigm); and (c) the step changes (signals) form a unidimensional "strength" continuum, i.e., their perceptual salience (detectability and/or subjective magnitude) increases with amplitude. The RT decrease is best documented for step increases in physical intensity (Pieron, 1920; Chocholle, 1940; Kohfeld, Santee, & Wallace, 1981a, b; Luce, 1986), but it has also been shown for non-intensity "strength" signals, such as velocity of a light

Correspondence and requests for reprints should be sent to Ehtibar N. Dzhafarov, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820.

0022-2496/92 $5.00

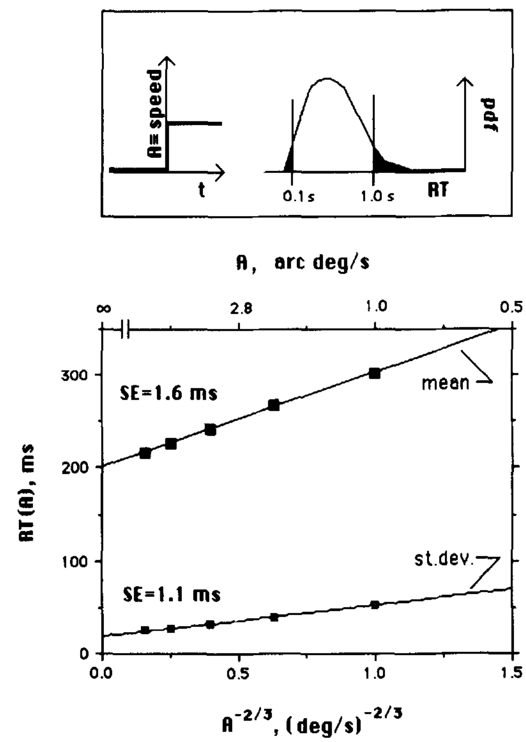

FIG. 1. Mean and standard deviation of the RT to visual motion onset (step-function signal with respect to motion speed; from an unpublished experiment by Dzhafarov, Sekuler, Allik, and Williams). The two statistics (computed over the RTs falling between 100 and 1000 ms) are arithmetically averaged over two subjects, which is justifiable because the expected dependence is linear with respect to theoretical parameters.

source (Ball & Sekuler, 1980; Tynan & Sekuler, 1982; Allik & Dzhafarov, 1984; Troscienko & Fahle, 1987; see Fig. 1).[1] The RT decrease is typically presented in terms of the mean and variance of (censored) RT distributions, but a more general fact can be stated: for a given percentile rank, RT percentile is a decreasing function of signal amplitude, except for the close-to-margins percentiles, which show weak or no ordinal dependence on amplitude.

A theoretical framework for this dependence is provided by the idea of a process evoked by signal onset, developing in time through successive stages, and terminating in an external response. Some stages notably (but not exclusively) motor ones, are not influenced by signal amplitude: their total duration can be termed the signal-independent component of RT. The complementary, signal-dependent, RT component decreases (stochastically) as a function of amplitude. There seems to be

[1] The linearizing transformation of velocity, $A^{-2/3}$, in Fig. 1, as well as the linearizing transformation of displacement, A^{-2}, in Figs. 5, 7, 9, and 14 have been derived from an elaborated version of a motion detection model proposed by Dzhafarov and Allik (Dzhafarov, 1982; Dzhafarov & Allik, 1984). The model is not discussed in this paper.

a concensus that these two RT components are mutually independent random variables. In the best available survey of the subject, Luce (1986, p. 97) expresses doubts that the independence hypothesis can be tested by psychophysical data. The independence, therefore, is assumed primarily because of its mathematical simplicity. However, unless distributional assumptions are made concerning the two RT components, the independence does not translate into simple algorithms for RT decomposition.

To illustrate some difficulties, consider a model in which a signal of amplitude A initiates a deterministic process, $\psi(t, A)$, developing in time, t. Detection occurs when this process reaches a random criterion level, $\mathbf{C}$. Then detection time is a random variable $\mathbf{D} = D(A, \mathbf{C})$, obtained by solving $\psi(t, A) = \mathbf{C}$ for t:[2] $\mathbf{D}$ is identified as the signal-dependent RT component, and RT is $D(A, \mathbf{C}) + \mathbf{F}$, where $\mathbf{F}$ denotes the total duration of signal-independent processes preceding and following the detection process. According to the independence assumption, we have $\kappa_i[\mathbf{RT}(A)] = \kappa_i[D(A, \mathbf{C})] + \kappa_i(\mathbf{F}]$ for cumulants κ_i of any order i.

For a typical psychophysical model, once it predicts the value of $D(A, \mathbf{C})$ for any given A and C, distributional assumptions about $\mathbf{C}$ and $\mathbf{F}$ are not essential. However, $\kappa_i[D(A, \mathbf{C})]$ cannot be found without such assumptions, and to make the model testable by RT data, the distribution of $\mathbf{C}$ will have to be derived from external considerations (such as the frequently misused invocation of the central limit theorem), or else it will have to be estimated from data. In the former case, a failure of the conjunction of the two hypotheses, about the function $D(A, \mathbf{C})$ and the distribution of $\mathbf{C}$, cannot be uniquely attributed to one of them—an undesirable situation if the distributional assumptions are not essential for the model. When the distribution is estimated from data, the model acquires too much freedom to be falsifiable by the observed $\kappa_i[\mathbf{RT}(A)]$, at least in a limited range of i. The situation improves only if $D(A, \mathbf{C})$ can be factorized into $f(\mathbf{C})\, D_0(A)$, in which case to fit the values of $\kappa_i[\mathbf{RT}(A)]$ all one has to do is to find coefficients $\kappa_i[f(\mathbf{C})]$ and $\kappa_i[\mathbf{F}]$.

Even in this case, however, the empirical applicability is limited to low-order cumulants only, typically the mean and the variance. High-order cumulants are difficult to estimate reliably, and they are sensitive to distribution tails and censoring procedures (Ashby & Townsend, 1980; Luce, 1986; Ratcliff, 1979). There are compelling reasons to believe that the lower tails of RT distributions contain false alarms mixed with signal-initiated responses, and the upper tails contain omission/distraction responses (see Section 1.5). As a result, RT distributions should be censored before their cumulants are computed. No censoring procedure, however, under the assumption of independence would separate "true" responses from signal-unrelated responses. In addition, any censoring procedure transforms the distribu-

[2] Here and throughout the text: (1) a symbol in boldface type, like $\mathbf{H}$, indicates a random variable; (2) the same symbol in regular type, like H, stands for a particular or arbitrary value of this variable; (3) the Pth percentile of a random variable $\mathbf{H}$ is denoted by H_P. RT, when in mathematical expressions, should always be read as a single symbol: $\mathbf{RT}$, RT_P, etc. Note that $\mathbf{H}(X)$ is a random variable depending on a parameter X, whereas $H(\mathbf{X})$ is a deterministic function of a random variable $\mathbf{X}$. All random variables are assumed to vary between but not within trials.

tions of **D** and **F** in an unknown way and introduces both an unknown dependence of **F** on A and an unknown interdependence between **D** and **F**. Consequently, the results of an analysis based on high-order RT cumulants (in fact beginning with the variance) will depend critically on the exact values, or percentile ranks, below and above which the RT distributions are trimmed. The same problems arise if the analysis is based on integral transforms of empirical RT distributions, rather than cumulants (see Section 1.3).

These difficulties would have to be taken as inherent limitations of the RT analysis of sensory processing if there were no reasonable and empirically testable alternatives to the assumption of independence. I show in this paper that this is not the case: an alternative is proposed and empirically justified that greatly simplifies algorithms of RT decomposition and considerably advances the analytical power of the RT paradigm. It is important to realize that mutual stochastic independence of the two RT components can in no way be inferred from the fact that only one of them, by definition, depends on signal amplitude: it is quite possible that both components depend on a common set of (signal-independent) random variables. In the decomposition $D(A, \mathbf{C}) + \mathbf{F}$, the variable **C** is A-independent, and it is logically possible that **F** and **C** are interdependent, inducing an interdependence between **F** and **D**.

In the simplest case, **F** might be a deterministic function of **C**, $\mathbf{F} = F(\mathbf{C})$, inducing a deterministic relation between **F** and **D**, for any given A. This possibility constitutes the essence of the RT decomposition model arrived at in this paper: the two RT components are deterministic (monotonic) functions of a single signal-independent random variable. In a sense, this is a conceptual opposite of the independence hypothesis. The decomposition model proposed is not based on parametric assumptions about underlying distributions, but it allows one to reconstruct these distributions a posteriori. The model also provides theoretical grounds for separating signal-initiated RTs from signal-unrelated responses.

Most of the theoretical predictions derived in this paper are related to the asymptotic behavior of RTs, as signal amplitude tends to its upper limit. Their empirical corroboration, therefore, requires sufficiently large amplitudes. At the same time, the model's applicability is only moderately limited: it is shown that at least for one class of signals all perfectly detectable amplitudes are "sufficiently large." The major advantage of the asymptotic analysis is that it does not require specific assumptions about sensory processes evoked by a given class of signals.

1.2. *Basic Assumptions*

Here I present two basic assumptions underlying all RT decomposition models considered in this paper. The discussion is confined to "true" reactions, evoked by signal onset (signal-unrelated RTs are discussed in Section 1.5).

1. *AdditiveDecomposition Assumption.* RT to a signal of amplitude A can be decomposed as (see footnote 2 for notation)

$$\mathbf{RT}(A) = D(A, \tilde{\mathbf{Z}}) + \mathbf{F}, \tag{1}$$

where

(a) $\mathbf{F} \geqslant 0$;

(b) $\tilde{\mathbf{Z}}$ is a set of random variables (not necessarily countable) such that the joint distribution of $(\mathbf{F}, \tilde{\mathbf{Z}})$ is independent of A;

(c) $D(A, \tilde{Z}) \geqslant 0$ is a strictly decreasing function of A for any value $\tilde{Z}$ of the random set $\tilde{\mathbf{Z}}$.[3]

$\mathbf{D}(A) = D(A, \tilde{\mathbf{Z}})$ is a signal-dependent component of RT. It is a random variable because it (deterministically) depends on the random set $\tilde{\mathbf{Z}}$. It is "stochastically decreasing" as A increases because it is strictly decreasing for any given $\tilde{Z}$ and the distribution of $\tilde{\mathbf{Z}}$ does not change with A. $\mathbf{F}$ is a signal-independent component of RT. The concept of "signal-independence" should not be taken as meaning "stimulus-independence," which is a much stronger term. It is only assumed that the distribution of $(\mathbf{F}, \tilde{\mathbf{Z}})$ does not vary as a function of amplitude of a specific stimulation parameter.

To appreciate the generality of Assumption 1, note that any random variable $\mathbf{D}(A)$ can be presented as a deterministic function D of A and of some random variables that are A-independent themselves (see Lemma 1.2.1 and Corollary 1.2.1 below). The assumption, therefore, does not impose any restrictions on the marginal distribution of $\mathbf{D}(A) = D(A, \tilde{\mathbf{Z}})$. Assumption 1 is only violated if for any choice of $\tilde{\mathbf{Z}}$ in Eq. (1), either the distribution of $(\mathbf{F}, \tilde{\mathbf{Z}})$ depends on A, or $D(A, \tilde{Z})$ does not decrease in A. The following lemma shows that this cannot happen in intuitively a very important case, when for a given value of $\mathbf{F}$ all percentiles of the conditional distribution of $\mathbf{D}(A)$ decrease with increasing A.

LEMMA 1.2.1.[4] *Let* $\mathbf{RT}(A) = \mathbf{D}(A) + \mathbf{F}$, *where* (1) $\mathbf{F} \geqslant 0$ *does not depend on* A; *and* (2) $D_P(A \mid F)$, *the* Pth *percentile of* $\mathbf{D}(A \mid \mathbf{F} = F) \geqslant 0$, *is a strictly decreasing function of* A *for any* P. *Then Assumption* 1 *is satisfied.*

COROLLARY 1.2.1. *Let* $\mathbf{RT}(A) = \mathbf{D}(A) + \mathbf{F}$, *where* (1) $\mathbf{F} \geqslant 0$ *does not depend on* A; (2) $D_P(A)$, *the* Pth *percentile of* $\mathbf{D}(A) \geqslant 0$, *is a strictly decreasing function of* A *for any* P; *and* (3) $\mathbf{D}(A)$ *and* $\mathbf{F}$ *are mutually independent. Then Assumption* 1 *is satisfied.*

This corollary shows that Assumption 1 is a direct generalization of the commonly accepted assumption of independence (see Section 1.1), provided that $\mathbf{D}(A)$ decreases in all percentiles with increasing A.

[3] Two technical comments: (a) the "value $\tilde{Z}$" means, of course, the set of values for all components of the random set $\tilde{\mathbf{Z}}$; (b) here and throughout the paper the terms "any," "always," and "all," when referring to values of random variables, should be taken with the usual possible exception of zero-probability sets.

[4] All proofs are presented in the Appendix.

2. *Asymptotic Differentiability Assumption.*[5] For any $\tilde{Z}$ (see footnote 3b), as $A \to \infty$, there asymptotically exists a nonzero finite derivative of $D(A, \tilde{Z})$ with respect to some positive function $s(A)$.

This means that $D(A, \tilde{Z})$ can be written as

$$D(A, \tilde{Z}) = D(\infty, \tilde{Z}) + C(\tilde{Z})\, s(A) + o\{s(A)\}, \tag{2}$$

where

(a) $s(A) > 0$;

(b) $C(\tilde{Z}) \neq 0$ is the finite asymptotic derivative $\partial D/\partial s$, $A \to \infty$;

(c) $D(\infty, \tilde{Z})$ is a non-negative limit function whose existence follows from the fact that $D(A, \tilde{Z})$ decreases while remaining nonnegative.

The difference

$$T(A, \tilde{Z}) = D(A, \tilde{Z}) - D(\infty, \tilde{Z}) = C(\tilde{Z})\, s(A) + o\{s(A)\}$$

is a strictly decreasing function of A converging to zero. The random variable $T(A, \tilde{Z})$ plays an important role in making the term "signal-dependence" unambiguous. As pointed out by Luce (1986, pp. 98–99), in general the decomposition of RT into signal-dependent and signal-independent components is not unique: if $\mathbf{D} + \mathbf{F}$ is such a decomposition, then so is $(\mathbf{D} + \mathbf{X}) + (-\mathbf{X} + \mathbf{F})$, provided $\mathbf{X}$ is signal-independent and the redefined components remain non-negative. $T(A, \tilde{\mathbf{Z}})$ is a "minimal" form of the signal-dependent RT component: it cannot be further additively decomposed. Correspondingly,

$$\mathbf{R} = D(\infty, \tilde{\mathbf{Z}}) + \mathbf{F}$$

is a "maximal" form of the signal-independent RT component.

Assumption 2 is rather unrestrictive. Its major purpose is to ensure asymptotic co-measurability of the central tendency of $T(A, \tilde{\mathbf{Z}})$ (median, mean) and the time-dimensioned variability of $T(A, \tilde{\mathbf{Z}})$ (interquartile range, standard deviation): as $A \to \infty$, the ratio of the two tends to a positive finite constant.[6]

The following simple results provide a detailed asymptotic characterization of the function $T(A, \tilde{Z})$, which will be used for theoretical analysis of RT distributions in Section 3.

[5] Without loss of generality, let A assume all values in $[0, \infty)$: for any other A_{sup} one can replace A by $(A^{-1} - A_{\text{sup}}^{-1})^{-1}$. Asymptotic statements in the following text are formulated as equations including $o\{s(A)\}$, which denotes a function of A such that $o\{s(A)\}/s(A) \to 0$ as $A \to \infty$. If $s(A)$ converges to zero, then $o\{s(A)\}$ converges to zero infinitely faster. In practice this means that $o\{s(A)\}$ can be neglected when s is sufficiently small (i.e., A is sufficiently large). $o\{1\}$ denotes any variable converging to zero.

[6] I am grateful to R. D. Luce (1991, personal communication) for subjecting Assumption 2 to a thorough analysis. He constructed a simple counter-example by putting $\tilde{\mathbf{Z}} = \mathbf{Z} > 0$, a single random variable, and $T(A, Z) = kA^{-Z}$. This yields $\mathbf{RT}(A) = \mathbf{R} + kA^{-\mathbf{Z}}$, for which, as proved by Luce, Assumption 2 is not satisfied. It is easy to show that as $A \to \infty$, the mean and median of $T(A, \mathbf{Z})$ are vanishing infinitely faster than its standard deviation and interquartile range. Intuitively, this is an "anomalous" behavior for a positive random variable, and Assumption 2 prevents this from happening.

LEMMA 1.2.2. *$T(A, \tilde{Z})$ is asymptotically factorizable into a product of a positive function $C(\tilde{Z})$ and a strictly decreasing positive function $s(A)$ vanishing at $A \to \infty$,*

$$T(A, \tilde{Z}) = C(\tilde{Z})\, s(A) + o\{s(A)\},$$

where $C(\tilde{Z})$ is unique, and $s(A)$ is asymptotically unique, except for positive scaling coefficients having reciprocal values for the two functions.

It immediately follows that

COROLLARY 1.2.2. *$T(A, \tilde{Z})$ asymptotically increases with $C(\tilde{Z})$, i.e.,* (a) *if $C(\tilde{Z}_1) \geqslant C(\tilde{Z}_2)$ then beginning with some value of A, $T(A, \tilde{Z}_1) \geqslant T(A, \tilde{Z}_2)$; and* (b) *if $C(\tilde{Z}_1) = C(\tilde{Z}_2)$ then $T(A, \tilde{Z}_1) - T(A, \tilde{Z}_2) = o\{s(A)\}$.*

The RT decomposition assumptions can now be summarized in the form

$$\mathbf{RT}(A) = T(A, \tilde{\mathbf{Z}}) + \mathbf{R}, \tag{3a}$$

$$\mathbf{R} = D(\infty, \tilde{\mathbf{Z}}) + \mathbf{F}, \tag{3b}$$

$$T(A, \tilde{\mathbf{Z}}) = C(\tilde{\mathbf{Z}})\, s(A) + o\{s(A)\}, \tag{3c}$$

where all terms and their relations are as above; obviously, $(\mathbf{R}, \tilde{\mathbf{Z}})$ is A-independent. Equation (3) will be referred to as the basic RT decomposition model (for signal-initiated responses). All models considered in this paper are particular cases or simple modifications of this model.

1.3. *RT Decomposition Models Based on the Assumption of Independence*

As mentioned in Section 1.1, there seems to be a consensus that the RT components $D(A, \tilde{\mathbf{Z}})$ and $\mathbf{F}$ in Eq. (1) are mutually independent random variables. In general this does not translate into mutual independence of $T(A, \tilde{\mathbf{Z}})$ and $\mathbf{R}$, the "minimal" signal-dependent component and the "maximal" signal-independent component, respectively. Indeed, $\mathbf{R}$ in Eq. (3) contains the component $D(\infty, \tilde{\mathbf{Z}})$ depending on $\tilde{\mathbf{Z}}$. If, however, the variability of $D(\infty, \tilde{\mathbf{Z}})$ happens to be much smaller than that of $\mathbf{F}$, then an approximate independence exists between $T(A, \tilde{\mathbf{Z}})$ and $\mathbf{R}$.

This suggests a strong version of the assumption of independence, in which $D(\infty, \tilde{\mathbf{Z}}) = \text{const}$, and the principal components of the basic RT decomposition, $T(A, \tilde{\mathbf{Z}})$ and $\mathbf{R}$, are mutually independent (Luce & Green, 1972; Kohfeld *et al.*, 1981a, b). Reflecting the two major properties of the model, the mutual independence of the RT components and the convergence to zero of the signal-dependent component, it will be referred to as the "asymptotic zero plus independence" ($AZ + I$) model.

The model is of special importance because it is the only independence-based model leading to distribution-free algorithms for RT decomposition. Indeed, as A

increases, the distribution function $\mathbb{F}_{\mathbf{RT}(A)}(t)$ tends to $\mathbb{F}_{R}(t)$.[7] Then $\mathbb{F}_{R}(t)$ can be estimated by $\mathbb{F}_{\mathbf{RT}(A^*)}(t)$ for a sufficiently large A^*, and the detection time distribution can be recovered from any $\mathbf{RT}(A)$ by deconvolving the estimated $\mathbb{f}_{R}(t)$ from $\mathbb{f}_{\mathbf{RT}(A)}(t)$. In terms of the moment generating functions (mgf),

$$\text{mgf}[T(A, \tilde{\mathbf{Z}})] \approx \text{mgf}[\mathbf{RT}(A)]/\text{mgf}[\mathbf{RT}(A^*)]. \tag{4}$$

The $AZ+I$ model will be shown to disagree with the observed asymptotic behavior of the RT distributions (Section 3.1). Due to its theoretical importance, however, it is worthwhile to see whether the predictions of the model can be improved by simple modifications while retaining the model's essential features. I consider two such modifications employing two schemata of mixing true RTs with signal-unrelated responses. One such schema is a 2-state mixture,

$$\mathbf{RT}(A) = \begin{cases} \mathbf{H} & \text{with probability } 1-\alpha; \\ T(A, \tilde{\mathbf{Z}}) + \mathbf{R} & \text{with probability } \alpha; \end{cases} \tag{5}$$

where $\mathbf{H}$ does not depend on A. Another schema assumes a race between signal-initiated and signal-unrelated processes. Denoting by $\mathbf{N}$ duration of the latter, the model states

$$\mathbf{RT}(A) = \min\{\mathbf{N}, T(A, \tilde{\mathbf{Z}})\} + \mathbf{R}. \tag{6}$$

$\mathbf{N}$ and $\tilde{\mathbf{Z}}$ are not necessarily mutually independent. As an example, consider again the model in which detection occurs if a deterministic process $\psi(t, A)$ reaches a variable criterion $\mathbf{C}$ (Section 1.1). In this model $\tilde{\mathbf{Z}} = \mathbf{C}$. Let an independent "noise process" $\mathbf{X}(t-\mathbf{t}_0)$ randomly launched at moments $\mathbf{t}_0$ race with $\psi(t, A)$ for the current criterion level, C. In this case $\mathbf{N}$ and $\mathbf{C}$ must positively covary. Note that both $\mathbf{N}$ in the race model and $\mathbf{H}$ in the 2-state model, being signal-unrelated, may attain non-positive values, counting from the signal onset.

It is shown (Section 3.3) that both modifications considered (which turn out to be asymptotically equivalent) can indeed improve the predictive power of the $AZ+I$ model, but only at the cost of assuming an unrealistically large proportion of signal-unrelated responses in RT distributions (50% or more for the RT data reported below).

1.4. *Overview of the Single-Variate RT Model*

A conventional approach, once the $AZ+I$ model is shown to fail, would be to seek the next simplest model as that in which $D(\infty, \tilde{\mathbf{Z}}) \neq \text{const.}$ i.e., its variability is not negligible, but $\tilde{\mathbf{Z}}$ and $\mathbf{F}$ in Eq. (3b) are still independent. The independence assumption then would be formally preserved, but not between the RT components

[7] Notation conventions: distribution and density functions of a random variable $\mathbf{H}$ at a value H will always be denoted by $\mathbb{F}_{H}(H)$ and $\mathbb{f}_{H}(H)$, respectively: the subscript indicates the name of the random variable. I will assume that $\mathbb{F}_{H}(H)$ for all time-dimensioned $\mathbf{H}$ are strictly increasing and differentiable on $(\inf H, \infty)$.

taken in the "minimal–maximal" form, $\mathbf{T}+\mathbf{R}$ (Eq. (3a)). It is shown, however, that such a model would be theoretically redundant in so far as the asymptotic RT distributions are concerned (Section 3.2). If $\mathbf{R}$ is comprised of elements depending on $\tilde{\mathbf{Z}}$ and elements independent of $\tilde{\mathbf{Z}}$, then the latter elements can be dropped from the model with no consequences for asymptotic predictions. Moreover, experimental data strongly suggest that the interdependence of $\mathbf{R}$ and $\tilde{\mathbf{Z}}$ has a particularly simple form: $\mathbf{R}$ is a deterministic increasing function of $T(A, \tilde{\mathbf{Z}})$ for any given A.

The simplest model incorporating this relationship is constructed by assuming that $T(A, \tilde{\mathbf{Z}})$ can be presented as $T(A, C(\tilde{\mathbf{Z}}))$ for all A, not only asymptotically, as it follows from Eq. (3). Related to Corollary 1.2.2, this means that $T(A, C(\tilde{\mathbf{Z}}))$ is a strictly increasing function of a single random variable, $\mathbf{C}=C(\tilde{\mathbf{Z}})$. The RT decomposition (Eq. (3)) then assumes the form

$$\mathbf{RT}(A)=T(A,\mathbf{C})+\mathbf{R}, \tag{7a}$$

$$\mathbf{R}=L(\mathbf{C})+m, \tag{7b}$$

$$T(A,\mathbf{C})=\mathbf{C}s(A)+o\{s(A)\}, \tag{7c}$$

where both $T(A, C)$ and $L(C)$ are increasing functions of C, and m is a positive constant (about 130 ms in the experiments reported). The constant $m=\inf\{R(C)\}$ is isolated because of its traditional interpretation as the "irreducible minimum": no signal-initiated RT can fall below this value.

Equation (7) will be referred to as the single-variate RT (SVRT) model: $\mathbf{RT}(A)$ in this model is a monotonic transformation of a single random variable, $\mathbf{C}$, rather than a combined effect of independent random variables selectively associated with A.[8] The function $L(C)$ is found to be a simple scale transformation, $L(\mathbf{C})=B\mathbf{C}$, where B is a positive constant. This finding simplifies RT analysis, but is not critical for the SVRT model.

The most important consequence of the SVRT model is that all RT percentiles of a given rank P, $\mathrm{RT}_P(A)$, correspond to a single value of the criterion C, namely, to the Pth percentile of $\mathbf{C}$, C_P. As a result, Eqs. (7a, b) can be rewritten as

$$\mathrm{RT}_P(A)=T(A, C_P)+L(C_P)+m. \tag{8}$$

Remarkable consequences of this equation are discussed in Section 1.6.

One interpretation of the random variable $\mathbf{C}$ in Eq. (7) is that it represents a preset "criterion" that has to be crossed by a signal-initiated deterministic process, $\psi(t, A)$, in order for detection to occur. This hypothetical process can always be constructed in such a way that solving the equation $\psi(t, A)=\mathbf{C}$ for t one gets $t=T(A,\mathbf{C})+L(\mathbf{C})=\mathbf{RT}(A)-m$. One can recognize here the variable criterion model proposed by Grice (1968, 1972; Grice, Nullmeyer, & Spiker, 1982) for simple RT. In a modified version, the process $\psi(t, A)$ can be constructed to yield

[8] Observe that if the SVRT model holds, then the only way to decompose $\mathbf{RT}(A)$ into a sum of mutually independent components, $\mathbf{D}(A)$ and $\mathbf{F}$, is to put $\mathbf{D}(A)=L(\mathbf{C})+T(A,\mathbf{C})$, and $\mathbf{F}=m$: a constant is simultaneously stochastically independent of and functionally related to any random variable.

$t = T(A, \mathbf{C})$ as a solution of $\psi(t, A) = \mathbf{C}$, so that $\mathbf{R} = L(\mathbf{C}) + m$ has to be added to this solution to get $\mathbf{RT}(A)$. Due to these and similar interpretations I refer to the random variable **C** as the "criterion." This term, however, should only be taken as a technical label. This paper does not address the question of what experimental manipulations may affect the distribution of **C**, and in what way: in particular, the "criterion" is not necessarily a "high-order" decision variable. To give an obvious example of another interpretation, consider $\mathbf{C}^{-1}\psi(t, A)$ to be a stochastic process which has to reach a fixed (unity) level for detection to occur: $\mathbf{C}^{-1}\psi(t, A) = 1$ (McGill, 1963). Since this equation is equivalent to $\psi(t, A) = \mathbf{C}$, the "criterion" may be a conceptual means to describe "sensory" variability of a particular kind, rather than a decision variable.[9]

1.5. *"True," "Too Fast," and "Too Slow" Reactions*

Equation (8) will be shown to hold in a wide percentile range, but not for very low and very high percentile ranks: as P approaches 0 or 100%, the dependence of $\mathrm{RT}_P(A)$ on A loses even ordinal regularity.

The problem with the upper-tail RT percentiles is to a large extent statistical: they cannot be reliably estimated from moderate-size samples due to their extreme positive skewness. Another factor is that with some probability the observer can be distracted and miss the signal onset. This would result in very long RTs (up to several seconds). Irrespective of A, these RTs will have percentile ranks exceeeding some relatively high value, say P^*. In the SVRT model the "distraction state" can be formally associated with an extremely high (or infinite) value of the criterion **C**, occurring with probability $1 - P^*$. As a result, Eq. (8) will not be affected at $P < P^*$.

The fact that the lower-tail RT percentiles can be close to zero, or even negative, indicates that the "true," signal-initiated responses are mixed here with signal-unrelated "false alarms." At the same time, the SVRT model for "true" reactions (Eqs. (7) and (8)) seems to account for all RT percentiles whose rank exceeds a relatively low value, P_* (unless $P > P^*$, as discussed above). A simple schema in which effectively all false alarms are concentrated below a certain RT percentile can be constructed as follows. Consider a "race" between $\psi(t, A)$ and $\mathbf{X}(t - \mathbf{t}_0)$ as discussed in Section 1.3 for the $AZ + I$ model (Eq. (6)). Assume that after either of the two processes reaches the current criterion level, C, the reaction continues as postulated in the SVRT model (Eq. 7(b)): the remaining time, $L(C) + m$, is an increasing function of C. It follows that the false alarm times are always shorter than the true RT for any given C. Assume now that the "noise process" is effectively bounded from above, i.e., $\mathrm{Prob}\{\mathbf{X}(t - \mathbf{t}_0) > C_*\}$ is negligibly small for some level C_*. Then Eq. (8) will hold for all P exceeding the percentile rank P_* of C_*.

1.6. *Operational Consequences of the SVRT Model*

The SVRT model greatly enhances one's ability to test models of sensory processing (with no a priori distributional assumptions) by means of the RT

[9] The present form of this paragraph is due to a most helpful discussion with R. D. Luce.

paradigm. Consider the model discussed in Section 1.1, predicting the detection time of a signal as a function of amplitude and criterion position, $D(A, C)$. To test these predictions, RT distributions are obtained for a number of amplitude values, $\{A_1, A_2, ..., A_n\}$. Let the percentiles $\mathrm{RT}_P(A_i)$ be computed across all these distributions, for values of P between some limits P_* and P^*. According to the SVRT model (Eq. (8)),

$$\mathrm{RT}_P(A_i) - \mathrm{RT}_P(A_j) = T(A_i, C_P) - T(A_j, C_p) = D(A_i, C_P) - D(A_j, C_P).$$

Fixing the value of P (say, at 50%) one can see that C_P is fixed across all values of A, so it can be estimated by means of a regression analysis (the form of the function D is specified by the model tested). Moving to another percentile level (say, 40 or 60%) one gets another test for the hypothetical function $D(A, C)$, and, as a by-product, an estimate of the 40th or 60th percentile of **C**. Continuing in this way, one gets as many tests of the model (within a single set of data) as there are percentile levels that can be reliably estimated from the RT distributions. Provided the model is corroborated between the percentile ranks P_* and P^*, as a by-product one reconstructs a posteriori the distribution of **C** between P_* and P^*. The situation is completely deterministic and, if the model is correct, the regression error is associated only with empirical estimation of the RT percentiles.

Moving downward from P_* and upward from P^* one can reach the percentile levels at which the fit deteriorates considerably. These levels would provide estimates of the percentile range containing only true RTs, separated from false alarms and omissions/distractions (Section 1.5). Note that ignoring marginal percentiles does not constitute a censoring (truncation) of the RT distributions: all the percentiles analyzed are determined from the entire distributions. At the same time, the relatively central percentiles are not affected by the values of the marginal ones—one of the well-known advantages of dealing with percentiles rather than moments or integral transforms (see Townsend & Ashby, 1983, pp. 95–98; Ratcliff, 1979).

If $D(A, C)$ is of interest in its "minimal" form only, $T(A, C)$, the analysis can be simplified further. From Eq. (7) we have $\lim \mathrm{RT}_P(A) = R_P$. The limit can be estimated by $\mathrm{RT}_P(A^*)$, where A^* is sufficiently large (the latter can be verified by observing that further increase in A leads to only negligible changes in $\mathrm{RT}_P(A)$—an "empirical form" of Cauchy's convergence criterion). Then

$$\mathrm{RT}_P(A) - \mathrm{RT}_P(A^*) \approx T(A, C_P), \tag{9}$$

which is a surprisingly simple analogue of the deconvolution discussed in Section 1.3 (Eq. (4)) and also more reliable as a numerical procedure.

2. Experimental Data

The theory presented in this paper is related to the results of two experiments on detection of instantaneous changes in spatial position of a small light source. The main experiment is a simple RT one in which the observers had to respond to a spatial shift as soon as it was noticed. To assess detectability of different spatial shifts, a Yes–No detection experiment was conducted in which the observers indicated whether they noticed or did not notice a shift. The experimental procedures were as follows.

2.1. *Method*

In both experiments the light source was a bright (>99% contrast) square-shaped spot ($1.1' \times 1.1'$) on a dark background, binocularly viewed from 220 cm (chin-rest, forehead support). A Macintosh Plus computer (60 Hz vertical refresh rate) presented the stimuli and recorded the subjects' responses. To prevent autokinetic motion, the room was illuminated dimly, allowing the display's borders ($4.8°\text{h} \times 3.2°\text{v}$) to be just discriminable.

A trial started when the subject pushed a designated "readiness" key on the computer keyboard with the left hand. This caused the spot to appear in the center of the display. The spot remained in this position for a random time interval (uniformly distributed between 1 and 2 s), adter which it disappeared and instantaneously reappeared in another position along the same horizontal line. The amplitude and direction of the displacement varied from trial to trial. The spot remained in the displayed position until the subject pushed a designated response key (one of two keys in the Yes–No experiment) on the keyboard with the right hand. Then the spot disappeared, ending the trial (response-terminated presentation mode).

Both experiments were carried out as a sequence of short blocks of trials, one or two blocks per subject per day. Within a block each amplitude value was presented 26 times, 13 times to the left and 13 times to the right. Except for this constraint, the presentation order was random.

Two right-handed subjects with corrected acuity participated in the simple RT experiment. The observers had to push the response key as quickly as possible as soon as a displacement was noticed, regardless of its direction. The displacement amplitudes in this experiment were 1.1′, 1.7′, 2.2′, 2.8′, 3.3′, 4.4′, 7.2′, and 19.9′. Observer SW participated in 9 blocks and observer RWS in 8 blocks (208 and 182 trials per amplitude, respectively).

Five different observers with normal or corrected acuity participated in the Yes–No detection experiment. The displacement amplitudes presented were 0.6′, 1.1′, and 1.7′, and, for two observers, 0′ (no displacement). The observers were asked to push the "Yes" key, if they noticed a displacement, "quickly enough" to indicate an "immediate impression." If no displacement was noticed within a sufficiently long time, counting from the trial start, the "No" key was pushed: all observers had been trained to wait at least 4–5 s to make sure. No feedback was given. Only type of response, "Yes" or "No," was recorded. Each observer participated in 4 blocks, 104 trials per amplitude per observer in total.

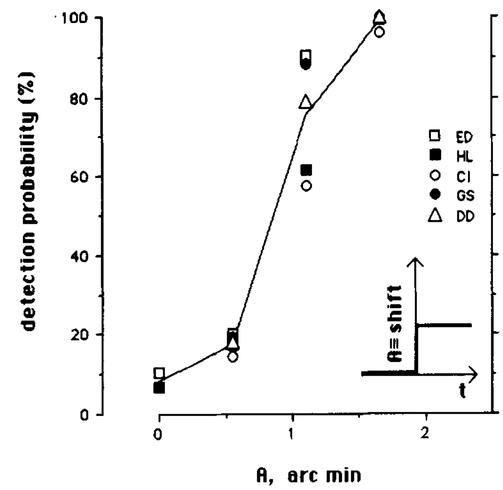

FIG. 2. Detection probability of instantaneous displacements (step-function signal with respect to spatial shift) of amplitudes 0, 0.55, 1.1, and 1.7 (arc min) for 5 observers; 104 responses per amplitude per observer; 0′ amplitude was presented to 2 observers only. The curve is drawn through across-subject means.

2.2. *Results*

(1) Distributions of the RTs to leftward and rightward displacements of the same amplitude were combined, as they were found sufficiently similar: the Smirnov-Kolmogorov two-sample statistics, $(n/2)^{1/2}D_{n,n}$, computed for 16 pairs of distributions (2 subjects by 8 amplitudes) range from 0.91 ($p \gg 0.2$) to 1.19 ($0.1 < p < 0.15$), with the median of 1.08 ($0.15 < p < 0.2$). The same similarity was found for the response frequencies in the Yes-No experiment (the χ^2 value pooled over all subjects and all amplitudes is 41.13, df = 34, p > 0.1). All data, therefore, are presented and analyzed as functions of amplitude only. The psychometric functions obtained in the Yes–No experiment are presented in Fig. 2. The empirical distribution functions obtained in the RT experiment are shown in Fig. 3.

(2) Figure 2 shows virtually perfect detectability of the 1.7′ displacements. Thus, of the amplitudes used in the RT experiment, all but the lowest, 1.1′, are virtually perfectly detectable. This "virtually" can be dropped for the next (after 1.7′) amplitude, 2.2′; the reason it was not included in the Yes–No experiment is that its perfect detectability was phenomenologically obvious. The detectability of the 1.1′ amplitude is imperfect and varies between 60 and 90%. Note in Fig. 3, RWS, an increase in the **RT**(1.1′) density at about the 70th percentile, indicating a bimodality. This bimodality is what can be expected if the upper percentiles of the distribution reflect trials when the displacement was not detected by this subject.[10]

[10] A more detailed analysis, not included in this paper, confirms that the detectability of the 1.1′ displacement by RWS was almost precisely 65%. The same amplitude was detected by subject SW about 90% of the time, and the ceiling effect prevents one from seeing a bimodality in this case.

(3) Figure 2 shows also about a 10% false alarm rate. Assuming the two experiments are roughly comparable, it might be expected then that signal-unrelated false alarms make up about 10% of RT trials (if the SVRT model holds, then the false alarm reactions should all be concentrated at the lower tails).

(4) In accordance with the general rule mentioned in Section 1.1, all percentiles of the RT distributions, $RT_P(A)$, are decreasing functions of A, provided P is

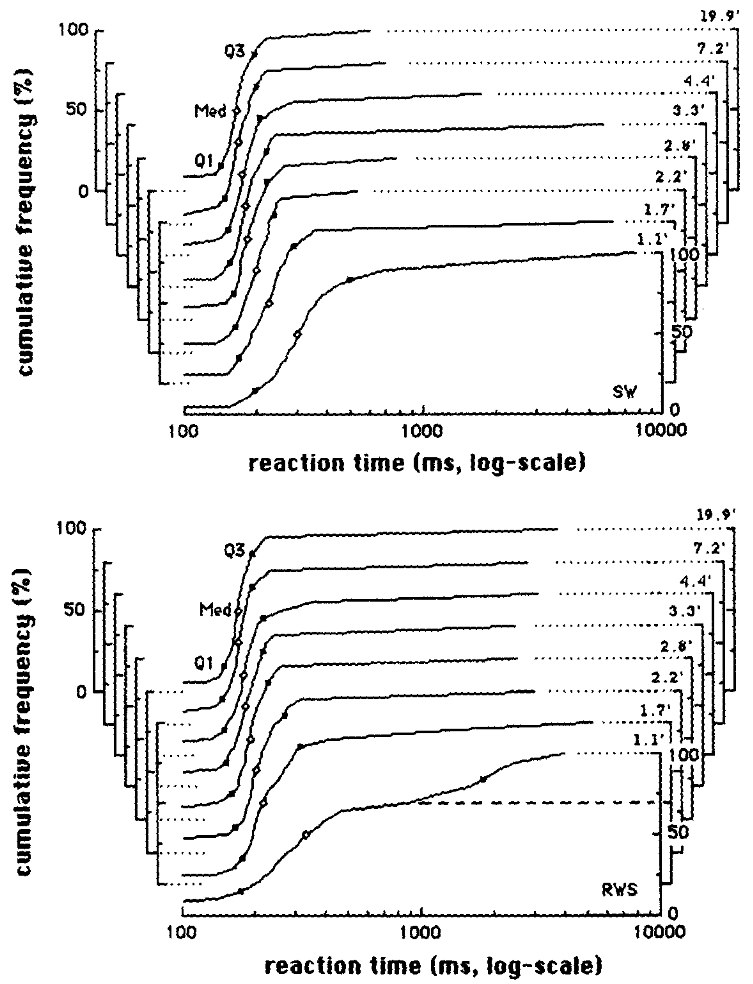

FIG. 3. Cumulative frequency distributions of the RT to instantaneous displacements of amplitudes 1.1, 1.7, 2.2, 2.8, 3.3, 4.4, 7.2, and 19.9 (arc min) for 2 observers; 208 RTs per amplitude for subject SW, 182 RTs per amplitude for subject RWS. The medians (open symbols) and two lateral quartiles (filled symbols) are indicated in each curve. For better discriminability the curves corresponding to different amplitudes are shifted by 20% steps with respect to each other. The amplitudes are shown above corresponding vertical axes. 100 and 0% levels for every curve are shown by dotted lines. The interrupted line in the panel for RWS, 1.1′, points out a density increase at about the 70% level, resulting in a second mode at about the 80–85% level.

not close to 0 or 100% (in our case, $10\text{–}15\% < P < 85\text{–}90\%$). This is demonstrated in Fig. 4. Note that the lower limit (10–15%) roughly agrees with the false alarm rate mentioned in paragraph (3).

(5) One can determine from Fig. 3 that the differences between corresponding percentiles of the RT distributions for $A = 19.9'$, the largest amplitude used, and 7.2′, the next largest amplitude, are very small, on the order of 1 ms. This means that as A exceeds 7.2′ $\mathbf{RT}(A)$ approaches its theoretical minimum, $\lim \mathbf{RT}(A) = \mathbf{R}$ (Eq. (3)).

2.3. *Asymptotically Linearizing Transformations*

Figure 5 shows that beginning with $A = 1.7'$, the three RT quartiles are well approximated by linear functions of A^{-2}. The linearity grossly deteriorates (especially for the third quartile) if the 1.1′ amplitude is included, and it improves, though very slightly, if the 1.7′ amplitude is excluded. This suggests that the transformation A^{-2} is asymptotically linearizing. Figure 6 shows that amplitude squaring is essential for the obtained linearity: the quality of fit for all three curves sharply deteriorates as the exponent shifts from -2 in either direction. The same results as in Figs. 5 and 6 have been obtained for other percentiles between the 15th and 85th (occasional small deviations from -2, as in the third quartile curve of Fig. 6, do not exhibit any systematic pattern). These percentiles are not shown here to avoid obscuring the graphs, but all are included in a more comprehensive analysis presented in Section 3.4.

It is a simple mathematical fact (see Theorem 3.1.2 below) that the mean RT, computed for the values of A and over the region of P where $\mathrm{RT}_P(A)$ is a linear function of A^{-2}, is itself a linear function of A^{-2}. Figure 7 demonstrates this fact

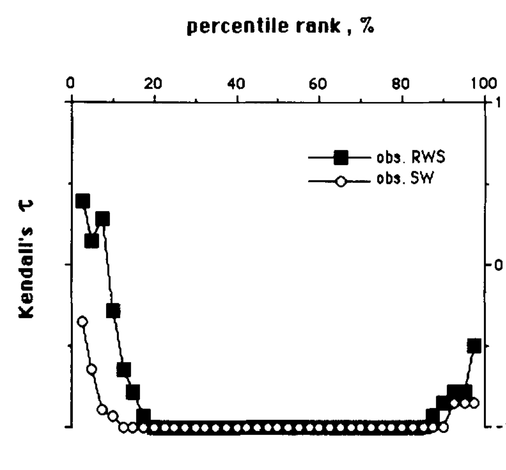

FIG. 4. Kendall's rank-order correlation between the amplitudes A and the percentiles $\mathrm{RT}_P(A)$ as a function of P (taken with 2.5% steps). Due to the extreme positive skewness of the distributions, the percentiles above the 90th could not be estimated reliably. The values of τ between -0.5 and 0.5 are not significantly different from 0 at 0.05 significance level.

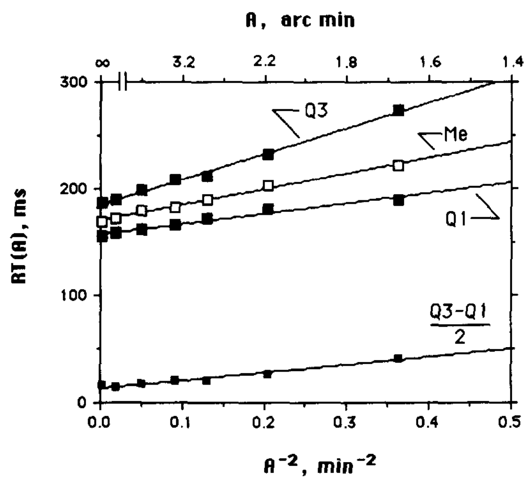

FIG. 5. RT quartiles and the semi-interquartile interval as functions of amplitude A ($A \geqslant 1.7$, i.e., $A = 1.1'$ is excluded). The statistics are arithmetically averaged over two subjects (RWS and SW), which is justified as in the legend of Fig. 1. The functions are linear with respect to A^{-2} (approximation error values can be read from Fig. 6, at $\beta = -2$).

for the RT distributions censored (truncated) below the 10th and above the 90th percentiles. The same is true for the RT standard deviation (not variance): it is linear in the same plots for the same censoring procedure. The linearity of the mean and standard deviation also holds for any other pair of cut-off percentile ranks taken between 10 and 90% (not necessarily symmetrically with respect to the median).

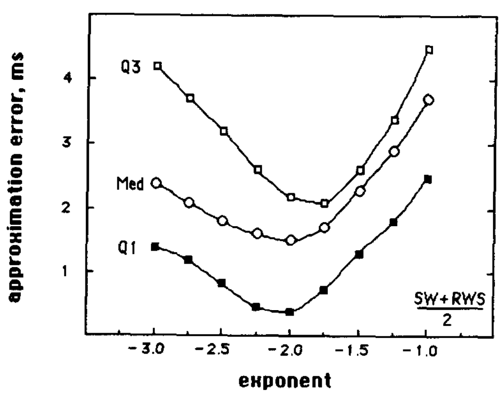

FIG. 6. Standard error of approximation of the three RT quartiles from Fig. 5 by linear functions of A^{β}, as a function of β.

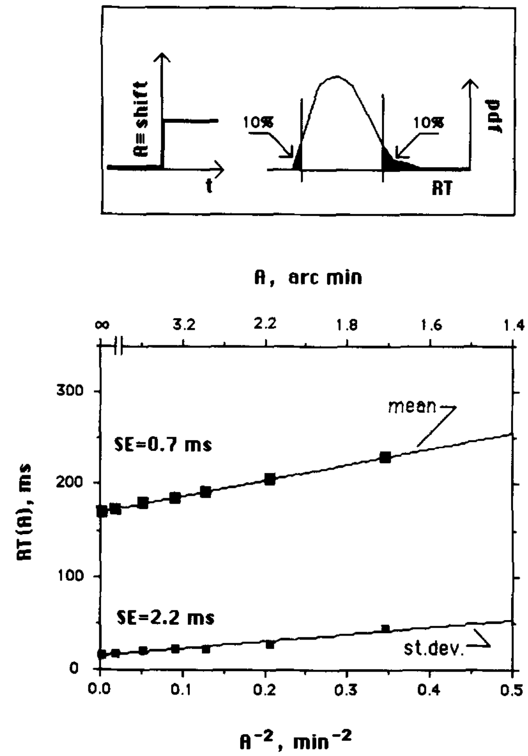

FIG. 7. Mean and standard deviation of the RT to instantaneous displacments ($A \geqslant 1.7'$), computed over the RTs falling between the 10th and 90th empirical percentiles, and arithmetically averaged over two subjects (RWS and SW; see legends of Figs. 1 and 5).

3. THEORETICAL ANALYSIS

3.1. *Dependence of the RT Percentiles on Signal Amplitude*

Figure 5 presents RT data in the format focal for this section: RTs of a given percentile rank P plotted against a transformation of A that linearizes the plots. It is shown in this section that such a linearizing transformation should exist for sufficiently large values of A if the basic RT model holds (Eq. (3)). Moreover, this transformation (asymptotically) equals the function $s(A)$ in the asymptotic factorization of $T(A, \tilde{Z})$ (Eq. (3c)).

The intercepts of the linearized $\mathrm{RT}_P(A)$ plots are theoretically equal to the corresponding percentiles R_P of the signal-independent RT component $\mathbf{R} = \mathbf{RT}(A) - T(A, \tilde{\mathbf{Z}})$. A theoretical meaning of the $\mathrm{RT}_P(A)$ slopes within the framework of the basic RT model is less obvious. The following theorem shows that these slopes equal $\mathbb{E}[\mathbf{C} \mid \mathbf{R} = R_p]$,[11] so the degree the slopes change with P reflects the degree of interdependence between $\mathbf{R}$ and the cirterion $\mathbf{C}$.

THEOREM 3.1.1. *In the basic RT decomposition model (Eq. (3))*

$$\mathrm{RT}_P(A) = R_P + \mathbb{E}[\mathbf{C} \mid \mathbf{R} = R_p]\, s(A) + o\{s(A)\}. \tag{10}$$

Three conclusions follow from this theorem when related to the RT_P vs $s(A)$ data. First, the function s in Eq. (3) can be empirically identified by finding the "best" linearizing transformation for sufficiently large values of A. In our case (Figs. 5 and 6), s asymptotically equals A^{-2}. Second, the range of "sufficiently large values of A" in the experiments reported coincides with the range of perfectly detectable signals. Finally, the data in Fig. 5 suggest that $\mathbf{R}$ and $\mathbf{C}$ have a considerable degree of positive covariation. Indeed, it immediately follows from Eq. 10 that

COROLLARY 3.1.1. *If* $\mathbf{R}$ *and* $\mathbf{C}$ *are mutually independent* ($AZ + I$ *model*), *then*

$$\mathrm{RT}_P = R_P + \mathbb{E}[\mathbf{C}]\, s(A) + o\{s(A)\}. \tag{11}$$

In other words, if $\mathbf{R}$ and $\mathbf{C}$ are mutually independent, the slopes of the linearized percentile curves have to be virtually identical for all P. This is obviously not the case: as P increases, both the intercepts and the slopes of the percentile curves increase (Fig. 5). Figure 8 demonstrates this for all analyzed percentiles; in fact, the slope of the linearized curves increases about 4–6 times faster than the intercept. The inset shows that this relationship is not critically determined by inclusion of the lower amplitudes, but is indeed an asymptotic property.

It is clear, therefore, that the $AZ + I$ model should be dismissed, even as a first approximation. At the same time the data are in conformity with the SVRT model. In this model only one value of $\mathbf{C}$ corresponds to $\mathbf{R} = R_P$, and this value is C_P (Eq. (8)). In terms of Eq. (10), $\mathbb{E}[\mathbf{C} \mid \mathbf{R} = R_P] = C_P$, which is obviously an increasing function of P (or of R_P).

A more traditional way to establish interdependence between $\mathbf{C}$ and $\mathbf{R}$ would be to linearize the mean and standard deviation of the RTs, rather than the percentiles. The following simple theorem shows that the asymptotically linearizing transformation s of A is the same in both cases.

THEOREM 3.1.2. *If the basic RT decomposition model holds for the RT distributions in a percentile rank region* (P_*, P^*), *then (see footnote* 11 *for notation)*

$$\mathbb{E}[\mathbf{RT}(A) \mid \mathsf{P}_A] = \mathbb{E}[\mathbf{R} \mid \mathsf{P}] + \mathbb{E}[\mathbf{Ç} \mid \mathsf{P}]s + o\{s\},$$

$$\mathbb{D}[\mathbf{RT}(A) \mid \mathsf{P}_A] = \mathbb{D}[\mathbf{R} \mid \mathsf{P}] + \mathbb{r}[\mathbf{Ç}, \mathbf{R} \mid \mathsf{P}]\, \mathbb{D}[\mathbf{Ç} \mid \mathsf{P}]s + o\{s\},$$

[11] Notation conventions: expected value and standard deviation (not variance) of a random variable $\mathbf{H}$ will be denoted by $\mathbb{E}(\mathbf{H})$ and $\mathbb{D}(\mathbf{H})$, respectively; correlation between $\mathbf{H}$ and $\mathbf{X}$ will be denoted by $\mathbb{r}(\mathbf{H}, \mathbf{X})$. If the distribution of $\mathbf{H}$ is conditioned on a certain statement P, the notation is $\mathbb{E}(\mathbf{H} \mid \mathsf{P})$ and $\mathbb{D}(\mathbf{H} \mid \mathsf{P})$. Symbols $\mathbb{E}(\mathbf{H} \mid X)$ and $\mathbb{D}(\mathbf{H} \mid X)$ can be used if the statement P is of the type $\mathbf{X} = X$. Finally, $\mathbb{E}(\mathbf{H} \mid \mathbf{X})$ is a random variable with values $\mathbb{E}(\mathbf{H} \mid X)$.

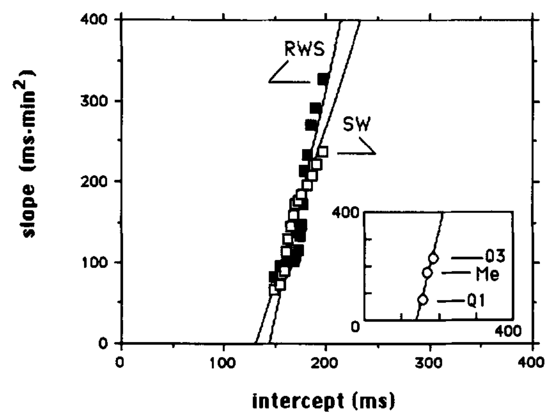

FIG. 8. Slopes and intercepts of the linear functions $b_0 + b_1 A^{-2}$ approximating $\mathrm{RT}_P(A)$, $A \geqslant 1.7$; symbols from left to right correspond to $P = 15, 20, ..., 80, 85\%$ (the linear functions for 25, 50, and 75% are shown in Fig. 5). The inset shows the slopes and intercepts of the quartile functions computed for three highest amplitudes only, 4.4′, 7.2′, and 19.9′ (averaged over two subjects).

where $\mathbf{P}_A$ *stands for* $\mathrm{RT}_{P_*}(A) \leqslant \mathbf{RT}(A) \leqslant \mathrm{RT}_{P^*}(A)$; $\mathbf{P}$ *stands for* $R_{P_*} \leqslant \mathbf{R} \leqslant R_{P^*}$; *and* $\mathbf{Ç} = \mathbb{E}[\mathbf{C} \mid \mathbf{R}]$.

As a result, if $\mathbf{C}$ and $\mathbf{R}$ are mutually independent, i.e., $\mathbb{r}[\mathbf{Ç}, \mathbf{R} \mid \mathbf{P}] = \mathbb{r}[\mathbf{C}, \mathbf{R} \mid \mathbf{P}] = \mathbb{r}[\mathbf{C}, \mathbf{R}] = 0$, then (omitting the conditions for simplicity)

$$\{\mathbb{D}[\mathbf{RT}(A)] - \mathbb{D}[\mathbf{R}]\} = o\{\mathbb{E}[\mathbf{RT}(A)] - \mathbb{E}[\mathbf{R}]\},$$

i.e., asymptotically $\mathbb{D}[\mathbf{RT}(A)]$ should be infinitely closer to a constant than $\mathbb{E}[\mathbf{RT}(A)]$. Figure 7 shows that this is not the case: $\mathbb{D}[\mathbf{RT}(A)] - \mathbb{D}[\mathbf{R}]$ is nearly proportional to $\mathbb{E}[\mathbf{RT}(A)] - \mathbb{E}[\mathbf{R}]$ ($\mathbb{E}[\mathbf{R}]$ and $\mathbb{D}[\mathbf{R}]$ are the intercepts of the straight lines).[12]

An analogous result is shown in Fig. 1 for the RT to uniform motion onset (the linearizing transformation here is $s = A^{-2/3}$, where A is velocity). Note, however, that it follows from the analysis just presented, that the cut-off points should be placed at fixed percentiles (as in Fig. 7) rather than at fixed RT values (as in Fig. 1). The similarity of the patterns displayed by the two figures should be attributed to the fact that the cut-off points in Fig. 1, 100 ms and 1000 ms, have roughly constant percentile ranks across the velocity values (at the same time, 100 ms is high enough to exclude most of the false alarms).

Summarizing, the experimental data are incompatible with the $AZ + I$ model, and strongly support the basic RT decomposition model with positive covariation

[12] All linear equations in this paper are explicitly formulated as asymptotic statements, hence in addition to sampling error the residuals in Figs. 5 and 7 (and later in Figs. 10–14) constain a systematic component induced by dropping the unknown $o\{s\}$-terms. Consequently, a formal goodness-of-fit analysis (when applicable), or tests of linearity against higher-order trends, are uninformative, aside from characterizing the statistical power of the tests.

of $\mathbf{T}(A)$ and $\mathbf{R}$. They are also in agreement with the SVRT model in which $\mathbf{T}(A)$ and $\mathbf{R}$ are deterministic increasing functions of each other for any given A. Although I am not aware of other experimental analyses where RT percentiles were asymptotically linearized, the approximately linear relationship between $\mathbb{D}[\mathbf{RT}(A)]$ and $\mathbb{E}[\mathbf{RT}(A)]$ for sufficiently large A seems to be a common finding (Cocholle, 1940; Green & Luce, 1971). Due to Theorem 3.1.2, this is an indication that the RT percentiles for other signal types also follow the SVRT model pattern.

3.2. *SVRT Model versus Basic RT Model*

Two models making identical predictions concerning $\mathrm{RT}_P(A)$ are operationally equivalent, whatever their formal differences. The following very simple theorem shows that this is the case with the basic RT model (Eq. (3)) and the SVRT model (Eq. (7)) in so far as the asymptotic RT distributions are concerned.

THEOREM 3.2.1. *If the basic RT decomposition model with $\mathbb{E}[\mathbf{C} \mid R_P]$ being an increasing function of P accounts for the asymptotic RT distributions in a percentile rank region (P_*, P^*), then in the same region the asymptotic RT distributions can be accounted for by the SVRT model with $\mathbf{C}_{\mathrm{SVRT}} = \mathbb{E}[\mathbf{C} \mid \mathbf{R}]_{\mathrm{basic}}$.*

The situation can now be viewed in the following way. The SVRT model and the $AZ+I$ model are two opposite and extreme cases of the basic RT model. The latter is corroborated by finding an asymptotically linearizing transformation s of A for the RT percentiles (a somewhat different approach is described in Section 3.4 below). Now, if the asymptotic slopes of the percentile curves do not change with their intercepts (equivalently, percentile ranks), then the basic RT model reduces to the $AZ+I$ model. If the slopes increase with percentile rank (which is the case for the experiments reported), then the basic RT model reduces to the SVRT model. Only if neither of the two situations took place (a decreasing or non-monotonic dependence), there would be a reason for keeping the basic RT model in its general form. This statement does not exclude the possibility that the SVRT model can be tested against the basic RT model using RT distributions for weak signals (see Section 4).

3.3. *Two-State Mixture And Race Models*

In this section I consider whether the compatibility of the independence hypothesis with the RT percentile data can be improved by the two modifications introduced in Section 1.3: the 2-state mixture model and the race model.

THEOREM 3.3.1. *Let the 2-state mixture model hold (Eq. 5), and let $\lambda(x)$ be the likelihood ratio function $\mathfrak{f}_H(x)/\mathfrak{f}_R(x)$. Then*

$$\mathrm{RT}_P(A) = R_P + \mathbf{C}(P)\,\mathbb{E}[\mathbf{C}]\,s + \mathrm{o}\{s\}, \tag{12}$$

where $\mathbf{C}(P) = [1 + \lambda(R_P)(1-\alpha)/\alpha]^{-1}$.

COROLLARY 3.3.1. *The asymptotic slope of the* RT_P *vs* $s(A)$ *curves increases with* P *if and only if the likelihood ratio* $\lambda(x)$ *is a monotonically decreasing function.*

One can see from this corollary that, unlike the $AZ+I$ model, its 2-state mixture modification can account for the fact that the slope of the linearized RT percentile curves increases with their intercept. One has to assume only that $\mathfrak{f}_H(x)/\mathfrak{f}_R(x)$ is a decreasing function. To understand why mixing signal-unrelated reactions with those generated according to the $AZ+I$ model makes such a change, note that Eq. (5) can be presented as

$$\mathbf{RT}(A) = \mathbf{bT}(A) + \{\mathbf{bR} + (1-\mathbf{b})\mathbf{H}\},$$

where $\mathbf{b}$ attains values 1 and 0 with probabilities α and $1-\alpha$, respectively. It is clear now that the signal-dependent component of $\mathbf{RT}$ is $\mathbf{bT}(A)$, rather than $\mathbf{T}(A)$ alone. Analogously, $\mathbf{bR}+(1-\mathbf{b})\mathbf{H}$ is the signal-independent component. The covariance of the two components equals

$$\mathbb{E}[\mathbf{b}^2\mathbf{TR}] - \mathbb{E}[\mathbf{bT}]\,\mathbb{E}[\mathbf{bR}+(1-\mathbf{b})\mathbf{H}] = \alpha(1-\alpha)\,\mathbb{E}[\mathbf{T}]\{\mathbb{E}[\mathbf{R}] - \mathbb{E}[\mathbf{H}]\}.$$

This value is positive, because, since $\lambda(x)$ is monotonically decreasing, $\mathbb{E}[\mathbf{R}] > \mathbb{E}[\mathbf{H}]$. In other words, the 2-state mixture improves the $AZ+I$ model's predictive power by de facto introducing a positive correlation between the signal-dependent and signal-independent RT components. In a sense, this defeats the original purpose: to retain the assumption of independence. Nevertheless, the model might still be considered a viable alternative to the SVRT model, if one could show that the estimated values of α and $\lambda(x)$ are reasonable. The parameter α can be roughly estimated by comparing the slopes of the linearized RT percentile curves to that of the linearized $\mathbb{E}[\mathbf{RT}]$-curve.

LEMMA 3.3.1. *If the* 2*-state mixture model holds, then for any* P,

$$\alpha \leqslant \frac{\text{asymptotic slope of } \mathbb{E}[\mathbf{RT}] \text{ vs } s(A)}{\text{asymptotic slope of } \mathrm{RT}_P \text{ vs } s(A)}. \tag{13}$$

Since $\lambda(R_P)$ is a decreasing function, the greater the P chosen, the closer the denominator of Eq. (13) approaches $\mathbb{E}[\mathbf{C}]$ and the more precise the estimate of α. However, one cannot use a percentile rank too close to 100%: for the considered data the 85th–90th percentile is about the highest for which the asymptotic linearization is still acceptable (Fig. 9, upper curve). The lower curves in Fig. 9 represent $\mathbb{E}[\mathbf{RT}]$.[13] The ratio of the slopes (Eq. (13)) shows that $\alpha < 0.55$ for both observers. Even assuming that at $P = 85$–90% $\lambda(R_P)$ is practically zero, so that

[13] It is not clear from the model how the RT distributions should be censored to compute $\mathbb{E}[\mathbf{RT}]$. Fortunately, $\mathbb{E}[\mathbf{RT}]$ is robust with respect to different types of censoring: as one can see from Fig. 9, the values of $\mathbb{E}[\mathbf{RT}]$ computed over data uncensored except for the exclusion of obvious upper-tail outliers are very close to $\mathbb{E}[\mathbf{RT}]$ computed as in Fig. 7.

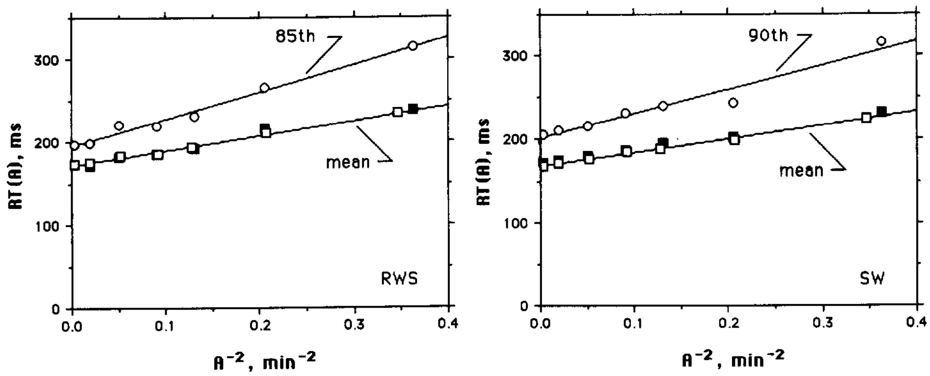

FIG. 9. Mean RT compared with a high-rank RT percentile plotted against A^{-2} ($A \geqslant 1.7$). The means computed from distributions censored as for Fig. 7 (open squares) coincide with the means computed from uncensored distributions (filled squares). The rank of the percentiles was chosen as the highest (in the series taken with 5% steps) for which the linear approximation is still acceptable. The ratio of the slopes is about 0.55 for both observers.

$\alpha \approx 0.55$, this means that almost half of all reactions are signal-unrelated. This clearly contradicts introspection and is an unrealistic possibility.

The situation with the second modification of the $AZ + I$ model introduced in Section 1.3, the race model, is exactly the same: the model can formally account for the positive slope-intercept covariation in the linearized RT percentile curves, but then half or more of all reactions should be signal-unrelated. The reason for this similarity is that the two models, 2-state and race, are asymptotically equivalent, as shown in the following theorem.

THEOREM 3.3.2. *The race model* (*Eq.* (6)) *is asymptotically equivalent to the 2-state mixture model* (*Eq.* (5)) *with*

$$\alpha_{2\text{-state}} = \text{Prob}\{\mathbf{N} > 0\}_{\text{race}}$$

$$\mathbb{F}_H(t)_{2\text{-state}} = \text{Prob}\{\mathbf{N} + \mathbf{R} < t \mid \mathbf{N} \leqslant 0\}_{\text{race}}$$

The asymptotic equivalence here means that for any X

$$\text{Prob}\{\mathbf{RT} < X\}_{\text{race}} = \text{Prob}\{\mathbf{RT} < X\}_{2\text{-state}} + o\{s\}.$$

In summary, unless one is ready to accept the possibility that half or more of all reactions in a simple RT experiment are signal-unrelated, one cannot accept the 2-state mixture and race models. In addition, these models do not account for the lack of a regular relationship between A and very low RT percentiles. It is formally possible, of course, to complement these models by the "fast false alarms" hypothesis presented in Section 1.5 for the SVRT model. This would amount, however, to adding still another type of signal-unrelated reaction to the already unreasonably large proportion.

3.4. *Percentile-versus-Percentile Curves*

In this section I describe a test for the SVRT model that does not require that one find the linearizing transformation $s(A)$ beforehand. I also show how to establish the form of the function $L(C)$ in Eq. (7), and how to estimate the constant m.

As a first step, recall that $\mathrm{RT}_P(A) - R_P = C_P s(A) + o(s)$ (see Theorem 3.2.1). Then for any two values A_1, A_2,

$$\mathrm{RT}_P(A_1) - R_P = (s_1/s_2)[\mathrm{RT}_P(A_2) - R_P] + o\{s_1\}.$$

Recall also that R_P can be estimated by $\mathrm{RT}_P(A^*)$, where A^* is a "very large amplitude," whose further increase does not lead to a non-negligible decrease in $\mathrm{RT}_P(A^*)$ (Eq. 9). If the SVRT model holds in a percentile range (P_*, P^*), then within this range one should expect that for sufficiently large A_1, A_2,

$$\mathrm{RT}_P(A_1) - \mathrm{RT}_P(A^*) \approx (s_1/s_2)[\mathrm{RT}_P(A_2) - \mathrm{RT}_P(A^*)]. \tag{14}$$

Plotted against each other, $\mathrm{RT}_P(A)$–$\mathrm{RT}_P(A^*)$ curves should form a fan of zero-intercept straight lines. Consider a class of amplitude values $\{A_1, A_2, ..., A_n\}$, in ascending order, with A_n taken for A^*. An economic way to show that Eq. (14) holds is to plot $\mathrm{RT}_P(A_i)$–$\mathrm{RT}_P(A^*)$, $i = 1, ..., n-1$, against their normalized mean,

$$\mu(P) = [\mathrm{RT}_P(\cdot) - \mathrm{RT}_P(A^*)]/[\mathrm{RT}.(\cdot) - \mathrm{RT}.(A^*)],$$

where the dot replacing a parameter indicates arithmetic averaging across all values of this parameter. If (and only if) Eq. (14) holds, then

$$\mathrm{RT}_P(A_i) - \mathrm{RT}_P(A^*) \approx b_i \mu(P), \tag{15}$$

where b_i does not depend on P and decreases with A. Figure 10 shows that Eq. (15) is satisfied with a reasonable precision for all perfectly detectable amplitudes between $P_* = 15\%$ and $P^* = 85\%$, with $A^* = 19.9'$ (Section 2.2(5)).

Note that $\mu(P)$ is an estimate of $C_P/C.$, hence the slopes b_i of the curves are estimates of $s_i C.$ measured in time units. Figure 11 confirms that the pattern in Fig. 10 is indeed only asymptotic: for the 1.1′ amplitude, whose detection probability is measurably less than 1, the approximation is very poor, even if one excludes the upper percentiles reflecting signal omission trials (Sections 1.5; 2.2(2)). Figure 11 shows also that zero-intercept straight lines are not a priori fittable to data in this format due to their ordinal-scale properties (see footnote 12).

The next problem to consider is the dependence of R_P on C_P. Since C_P (in units of $C.$) is estimated by $\mu(P)$, and R_P is estimated by $\mathrm{RT}_P(A^*)$, the simplest way to deal with the problem is to plot $\mathrm{RT}_P(A^*)$ against $\mu(P)$. To make the analysis more convincing, one can look at the dependence on $\mu(P)$ of $\mathrm{RT}_P(A_i)$ for all values of A shown in Fig. 10, in addition to A^* (Fig. 12).

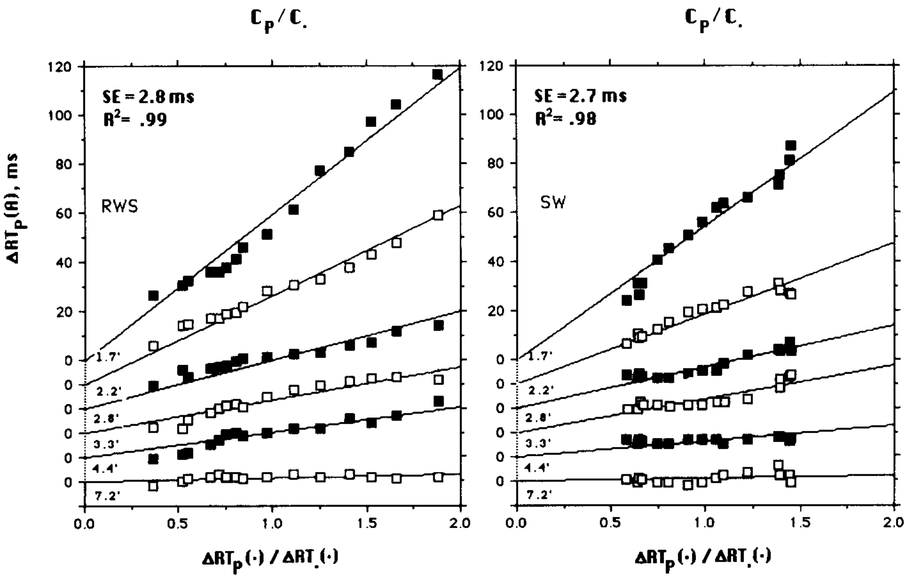

FIG. 10. Empirical estimates of $\Delta RT_P(A) = RT_P(A) - \lim RT_P(A)$, $A \geqslant 1.7$, plotted against the means of these estimates across A, $\Delta RT_P(\cdot)$, normalized by their grand mean, $\Delta RT.(\cdot)$; $\lim RT_P(A)$ is estimated by $RT_P(A = 19.9')$; $P = 15, 20, ..., 80, 85\%$. Theoretically, the graphs represent estimates of sC_P plotted against estimates of $C_P/C.$.

Surprisingly, the $RT_P(A)$ vs $\mu(P)$ functions turn out to be linear with a common intercept, indicating that R_P is a linear function of C_P:

$$R_P = (BC.)(C_P/C.) + m = BC_P + m. \tag{16}$$

The "irreducible minimum" m in Fig. 12 is obtained by linear extrapolation, even though the percentiles close to sero have been excluded from the analysis. There is no contradiction here, because the low-tail percentiles reflect a mixture of true reactions and false alarms, as discussed in Section 1.5. The constancy of the intercept across different curves can be considered a validation of this hypothesis.

Figure 13 presents a different way to establish that R_P is a linear function of C_P, and to estimate m. Since $RT_P(A) - R_P = Cs + o(s)$, it follows from Eq. (16) that

$$RT_P(A) = (s + B)\, C_P + m + o\{s\}. \tag{17}$$

Then for any two sufficiently large values from $\{A_1, A_2, ..., A_n\}$,

$$RT_P(A_i) \approx b_{ij} RT_P(A_j) + m(1 - b_{ij}),$$

where b_{ij} stands for $(s_i + B)/(s_j + B)$. If one computes all pairwise regression lines $RT_P(A_i)$ vs $RT_P(A_j)$, and plots intercept values against the values of (1-slope), the relation should be described by a zero-intercept linear function. The slope of this

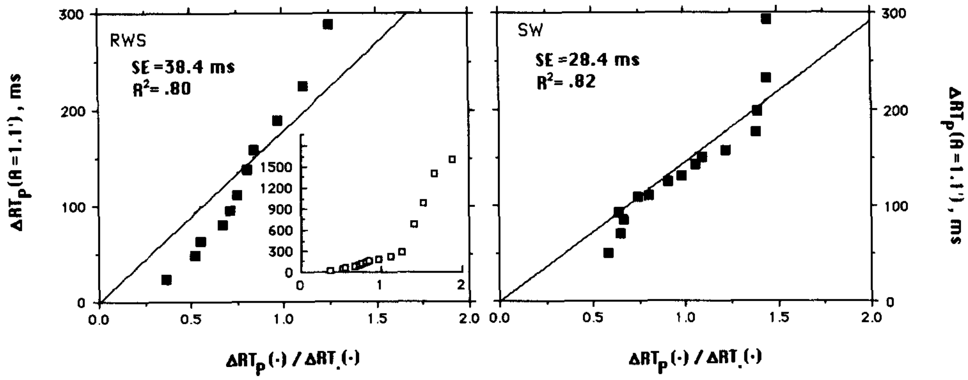

FIG. 11. Same as in the Fig. 10, but $A = 1.1'$. The graph for RWS does not include the four higher percentiles (70th through 85th), because of their very large values (all percentiles are shown in the inset).

linear function is an estimator of m. Figure 13 shows that this prediction is correct, and the estimates of m practically coincide with those obtained from Fig 12.

Still another estimate for m can be derived from Fig. 7, by plotting $\mathbb{E}[\mathbf{RT}(A)]$ against $\mathbb{D}[\mathbf{RT}(A)]$ (omitting for simplicity the censoring condition $\mathbf{P}_A$). Indeed, it follows from Eq. (17), that $\mathbb{E}[\mathbf{RT}(A)]$ must be a linear function of $\mathbb{D}[\mathbf{RT}(A)]$, with

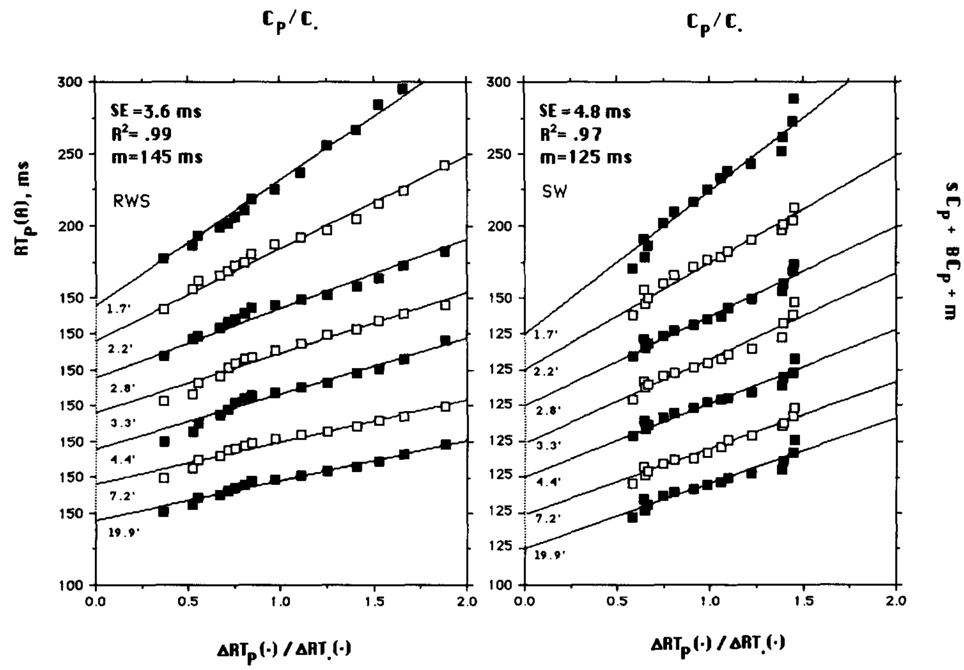

FIG. 12. $\mathrm{RT}_P(A)$, $A \geqslant 1.7$, plotted against the same horizontal axis as in Figs. 10 and 11); $P = (15, 20, ..., 80, 85\%$. Theoretically, the vertical axis represents estimates of $sC_P + BC_P + m$.

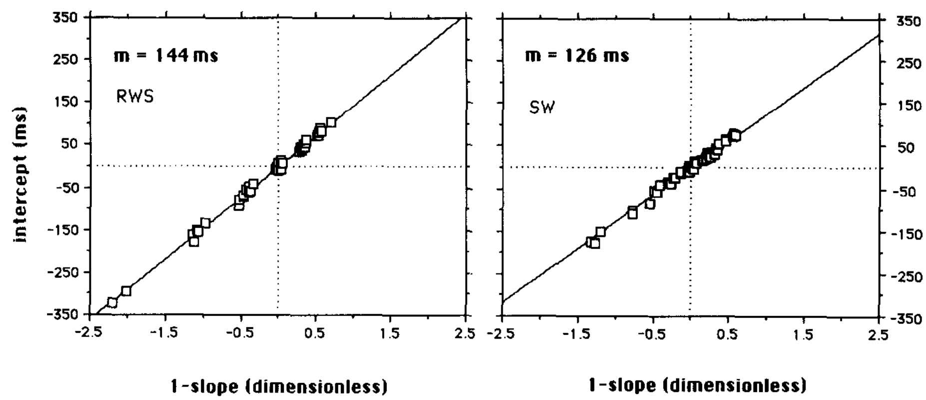

FIG. 13. The values of b_0 and $1-b_1$ of the linear functions $\mathrm{RT}_P(A_j)=b_0+b_1\mathrm{RT}_P(A_i)$, computed for all amplitude pairs (A_i, A_j) chosen from the set $A \geqslant 1.7'$; $P=15, 20, \ldots, 80, 85\%$. The dotted lines separate the regression coefficients for $A_j > A_i$ (upper right quadrant) from those for $A_j < A_i$ (lower left quadrant). The slope of the straight line fitted to the $[b_0, 1-b_1]$ pairs is an estimate of parameter m.

the intercept estimating m. (The computations yield 127 ms for SW and 146 ms for RWS, in agreement with Figs. 12 and 13).

I return now to the linearizing transformation $s(A)$. Its identification was not necessary so far in this section, and the SVRT model was corroborated without regard to the form of this transformation. Once established, however, the model allows one to solve the identification problem as a separate issue. The procedure is analogous to that used in Section 2.3, Figs. 5 and 6, but this time it includes all RT percentiles considered, and does so in a more economic way.

The slopes of the linear functions in Figs. 10 and 12 are estimators of the theoretical values $sC.$ and $sC. + BC.$, respectively. Plotted against any function of A, these slope values would form two parallel curves, with the vertical separation equal to $BC.$. To find the transformation $s(A)$ means to linearize these two parallel curves.

In agreement with the previous analysis, Fig. 14 shows that putting $s = A^{-2}$ yields an almost perfect linear fit. Figure 15 confirms that if the transformation is sought in the form A^{β}, $\beta < 0$, the best linear fit is indeed obtained at β about -2.

4. CONCLUSION

I have shown in this paper that the problem of interdependence between the signal-dependent and signal-independent RT components can be meaningfully addressed by psychophysical means, based on plausible initial assumptions. The two RT components have been shown to positively covary, which disagrees with the commonly accepted independence hypothesis. The SVRT model proposed is the simplest conceptual opposite of the independence hypothesis and the simplest variant of the basic RT decomposition model that agrees with data. In this model

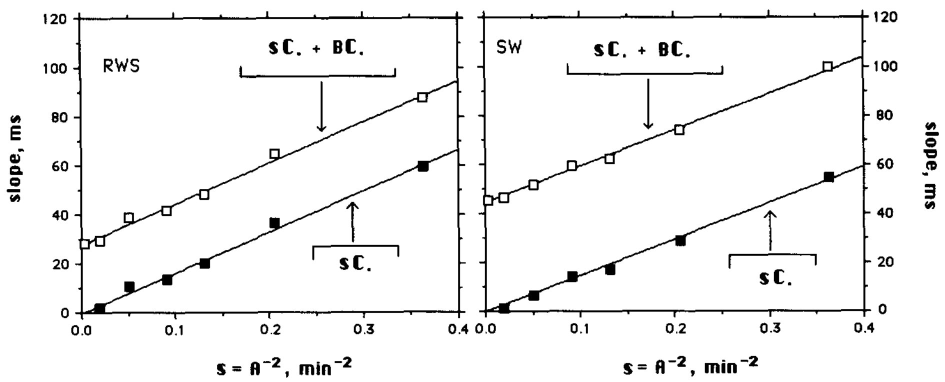

FIG. 14. The slopes of the linear functions from Fig. 10 (filled symbols) and Fig. 12 (open symbols), plotted against $s = A^{-2}$. Theoretically, the two functions are, respectively, $sC.$ and $sC. + BC.$, i.e., they are linear with the same slope, C. (approximation error and determination coefficient values can be read from Fig. 15, at $\beta = -2$).

the additive RT components are monotonic functions of each other for any given signal amplitude, and their variability is accounted for by a single random variable, termed the "criterion."

One way to describe the SVRT model is to say that the "criterion" **C** is a "double-readiness" factor: readiness to detect and readiness to respond. The higher the criterion the lower the readiness both to detect and to respond; hence, the components $\mathbf{T}(A) = T(A, \mathbf{C})$ and $\mathbf{R} = L(\mathbf{C}) + m$ increase or decrease together for any given A. This interpretation is especially plausible if one accepts Grice's modelling schema (see Section 1.4, last paragraph). As shown in the same paragraph, **C** can also be interpreted as a factor controlling sensory accumulation rate, a multiplicative perturbation of a deterministic accumulation process, $\psi(t, A)$, by a signal-independent noise. The fact that **C** is signal-independent suggests that the source of this noise is relatively central: a transient "inhibition" (attenuation) state is imposed

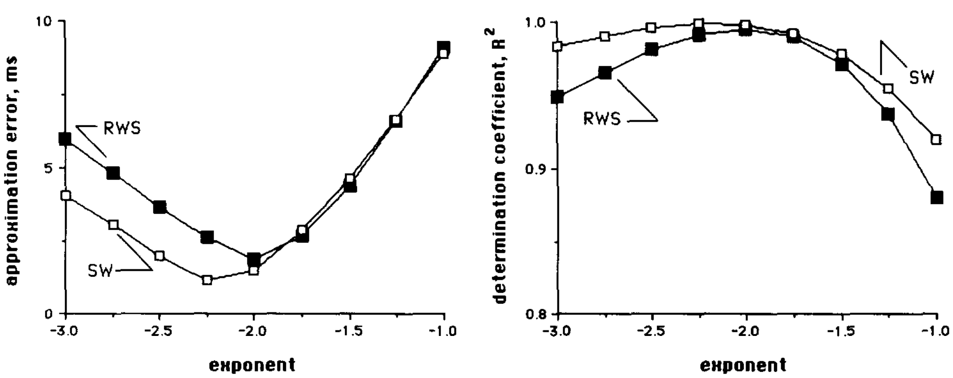

FIG. 15. Standard error of approximation and determination coefficient obtained when A^{-2} in Fig. 14 is replaced by other powers of A.

on an entire network of signal-specific sensory encoders (e.g., Reichardt-type bilocal correlators in the case of motion/displacement detection; Reichardt, 1961; see Nakayama, 1985, for a review). The SVRT model then can be viewed as suggesting that this transient inhibition affects activation transmission through all processing subcomponents, signal-dependent or not, possibly including the activation transmission from sensory encoders to motor units.

It remains to be seen whether the SVRT model applies to weak signals, those whose detection probability is measurably less than 1. If Assumption 1 (Section 1.2) holds, the distribution of **R** extracted from asymptotic RT distributions should remain the same for all signal amplitudes, however small; otherwise, the joint distribution of ($\tilde{\mathbf{Z}}$, **R**) would be signal-dependent. In particular , the approximation of **R** by $L(C[\tilde{\mathbf{Z}})) + m$ should apply to all amplitudes. However, the approximation of $T(A, \tilde{\mathbf{Z}})$ by $T(A, C(\tilde{\mathbf{Z}}))$, an increasing function of $C(\tilde{\mathbf{Z}})$, may not apply to small amplitudes. It is possible, therefore, that even though **R** is the same, its extraction from weak-signal RTs requires different decomposition algorithms. It is also possible, of course, that the Assumption 1 itself holds only asymptotically. Thus, it is not certain that by subtracting an asymptotic estimate of R_P from $\mathrm{RT}_P(A)$ for small A, one would obtain an estimate of $T_P(A)$. This prediction, however, can be empirically tested in conjunction with a specific model of sensory processing (see Section 1.6).

It is possible that further analyses will demonstrate limitations of the SVRT model as a general model of simple reaction time, either for small signal values or even as an asymptotic approximation. However, due to its remarkable simplicity and openness to empirical falsification, the model can be viewed as a new normative schema, in addition to the traditional schema based on the independence hypothesis.

APPENDIX: Proofs

Proof of Lemma 1.2.1. $\mathbf{D}(A)$ can always be presented as $D(A, \mathbf{F}, \mathbf{P})$, a deterministic function of A, **F**, and a random variable **P** uniformly distributed between 0 and 100. Indeed, put $D(A, F, P) = D_P(A \mid F)$, the Pth percentile of the conditional distribution of $\mathbf{D}(A \mid \mathbf{F} = F)$, $0 \leqslant P \leqslant 100$. **P** and **F** are mutually independent and independent of A. Hence the joint distribution of (**F**, **P**) does not depend on A, and neither does the joint distribution of (**F**, **P**, **F**). The proof is completed by putting $\tilde{\mathbf{Z}} = (\mathbf{F}, \mathbf{P})$, and observing that (a), (b), and (c) in Assumption 1 are satisfied.

In particular, if $\mathbf{D}(A)$ and **F** are mutually independent, then $D(A, F, P) = D(A, P) = D_P(A)$, the Pth percentile of $\mathbf{D}(A) = D(A, \mathbf{P})$. Putting $\tilde{\mathbf{Z}} = \mathbf{P}$, we get a proof of Corollary 1.2.1. ▮

Proof of Lemma 1.2.2 (omitting arguments of functions C and s for short). That C is positive and s is vanishing at $A \to \infty$ follows trivially from Eq. (2). The uniqueness properties are proved by equating $C_1 s_1 + o\{s_1\}$ and $C_2 s_2 + o(s_2\}$, and observing that then $s_1/s_2 \to C_2/C_1$. Since the limit of s_1/s_2 does not depend on $\tilde{Z}$, and C_2/C_1 does not change with A, this asymptotic equality can only be satisfied if $\lambda = C_2/C_1$ is a constant, and $s_1/s_2 \to \lambda$. Then $C_1 = C_2/\lambda$, and $s_1 = \lambda s_2 + o\{s_2\}$. Finally, to prove that $s(A)$ can be chosen to be strictly decreasing, put $s(A) = T(A, \tilde{Z}_0)$ for a fixed value $\tilde{Z}_0$ of the random set $\tilde{\mathbf{Z}}$. ∎

Proof of Theorem 3.1.1.[14] The percentiles RT_P and R_P are defined from

$$\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P\} = P; \tag{*1}$$

$$\mathrm{Prob}\{\mathbf{R} < R_P\} = P. \tag{*2}$$

$\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P\} = \mathrm{Prob}\{\mathbf{R} < \mathrm{RT}_P - \mathbf{T}\}$ can be presented in terms of the conditional distribution functions $\mathbb{F}_{R|T}(R|T)$ as

$$\int_0^{\mathrm{RT}_P} \mathbb{F}_{R|T}(\mathrm{RT}_P - T | T)\, d\mathbb{F}_T(T) = P. \tag{*3}$$

In the region $RT_P - T > 0$ $\mathbb{F}_{R|T}(\mathrm{RT}_P - T | T)$ can be expanded as

$$\mathbb{F}_{R|T}(\mathrm{RT}_P | T) - \mathbb{f}_{R|T}(\mathrm{RT}_P | T) T + o\{T\}.$$

Observing that both $o\{T\}$ and $\int_0^{\mathrm{RT}_P} o\{T\}\, d\mathbb{F}_T(T)$ are $o\{s\}$, one can rewrite (*3) as

$$\int_0^{\mathrm{RT}_P} \mathbb{F}_{R|T}(\mathrm{RT}_P | T)\, d\mathbb{F}_T(T) - \int_0^{\mathrm{RT}_P} T \mathbb{f}_{R|T}(\mathrm{RT}_P | T)\, d\mathbb{F}_T(T) + o\{s\} = P. \tag{*4}$$

In the second integral $\mathbb{f}_{R|T}(\mathrm{RT}_P | T)$ can be replaced with $\mathbb{f}_{R|T}(R_P | T)$ because

$$\int_0^{\mathrm{RT}_P} T[\mathbb{f}_{R|T}(\mathrm{RT}_P | T) - \mathbb{f}_{R|T}(R_P | T)]\, d\mathbb{F}_T(T) = o\{s\}.$$

Indeed, the expression in the brackets is $o\{1\}$, so the integrated function is $o\{T\}$. After this substitution the second integral in (*4) can be rewritten as

$$\int_0^{\mathrm{RT}_P} T \mathbb{f}_{R|T}(R_P | T)\, d\mathbb{f}_T(T) = \mathbb{f}_R(R_P) \int_0^{\mathrm{RT}_P} T\, d\mathbb{F}_{T|R}(T | R_P), \tag{*5}$$

because

$$\mathbb{f}_{R|T}(R_P | T)\, d\mathbb{F}_T(T) = \mathbb{f}_R(R_P)\, d\mathbb{f}_{T|R}(T | R_P) = \mathbb{f}_{R,T}(R_P, T)\, dT.$$

Next, the upper integration limits for both integrals in (*4) can be replaced with ∞. Indeed, for the first integral[15]

$$\int_{\mathrm{RT}_P}^{\infty} \mathbb{F}_{R|T}(\mathrm{RT}_P | T)\, d\mathbb{F}_T(T) < \int_{\mathrm{RT}_P}^{\infty} d\mathbb{F}_T(T) < \mathbb{E}[\mathbf{T}^2]/\mathrm{RT}_P^2,$$

[14] Notation conventions: (1) in this and following theorems, to make mathematical expressions less cumbersome, I will omit the argument A when referring to random variables or functions depending on A: thus, s, **RT**, RT_P, **T**, etc., will stand for $s(A)$, $\mathbf{RT}(A)$, $\mathrm{RT}_P(A)$, $\mathbf{T}(A)$, etc.; (2) conditional distribution and density functions of a random variable **H** given a certain value of a variable **G** will be denoted by $\mathbb{F}_{H|G}(H|G)$ and $\mathbb{f}_{H|G}(H|G)$, respectively.

[15] I use here Chebyshev's inequality for non-negative variables: $1 - \mathbb{F}_X(x) < \mathbb{E}[\mathbf{X}^2]/x^2$.

and since $\mathbb{E}[\mathbf{T}^2] = o\{s\}$ whereas $\mathrm{RT}_P^2 \to R_P^2$, their ratio is $o\{s\}$. For the second integral, taken in the form (*5),[16]

$$\mathbb{f}_R(R_P) \int_{\mathrm{RT}_P}^{\infty} T\, d\mathbb{F}_{T|R}(T \mid R_P) < \mathbb{f}_R(R_P)\, \mathbb{E}[\mathbf{T}^2 \mid R_P]/\mathrm{RT}_P,$$

and, again, $\mathbb{E}[\mathbf{T}^2 \mid R_P] = o\{s\}$. Now (*4) can be presented as

$$\int_0^{\infty} \mathbb{F}_{R|T}(\mathrm{RT}_P \mid T)\, d\mathbb{F}_T(T) - \mathbb{f}_R(R_P) \int_0^{\infty} T\, d\mathbb{F}_{T|R}(T \mid R_P) + o\{s\} = P. \qquad (*6)$$

At the same time, due to (*2),

$$\int_0^{\infty} \mathbb{F}_{R|T}(R_P \mid T)\, d\mathbb{F}_T(T) = P. \qquad (*7)$$

Subtracting (*7) from (*6), and observing that

$$\mathbb{F}_{R|T}(\mathrm{RT}_P \mid T) - \mathbb{F}_{R|T}(R_P \mid T) = \mathbb{f}_{R|T}(R_P \mid T)(\mathrm{RT}_P - R_P) + o\{\mathrm{RT}_P - R_P\},$$

one gets the following equation:

$$(\mathrm{RT}_P - R_P) \int_0^{\infty} \mathbb{f}_{R|T}(R_P \mid T)\, d\mathbb{F}_T(T) + o\{RT_P - R_P\}$$
$$= \mathbb{f}_R(R_P) \int_0^{\infty} T\, d\mathbb{F}_{T|R}(T \mid R_P) + o\{s\}. \qquad (*8)$$

The expression $o\{\mathrm{RT}_P - R_P\}$ can be dropped, because it is $o\{s\}$, as one can see by dividing both sides of the equation by s. Using the same transformation as in (*5), the left-hand integral in (*8) can be presented as

$$\int_0^{\infty} \mathbb{f}_{R|T}(R_P \mid T)\, d\mathbb{F}_T(T) = \mathbb{f}_R(R_P) \int_0^{\infty} d\mathbb{F}_{T|R}(T \mid R_P) = \mathbb{f}_R(R_P).$$

Since the right-hand integral in (*8) represents $\mathbb{E}[\mathbf{T} \mid R_P]$, the equation assumes the form

$$\mathbb{f}_R(R_P)(\mathrm{RT}_P - R_P) = \mathbb{f}_R(R_P)\, \mathbb{E}[\mathbf{T} \mid R_P] + o\{s\}.$$

Dividing both sides by $\mathbb{f}_R(R_P)$, and observing that

$$\mathbb{E}[\mathbf{T} \mid R_P] = \mathbb{E}[\mathbf{C} \mid R_P]\, s + o\{s\},$$

one comes to the statement of the theorem. ∎

[16] Here I use an inequality for non-negative variables closely related to Chebyshev's: $\mathbb{E}[\mathbf{X} \mid \mathbf{X} > x] \times (1 - \mathbb{F}_X(x)) < \mathbb{E}[\mathbf{X}^2]/x$. Its proof is analogous to that of Chebyshev's inequality.

Proof of Theorem 3.1.2. Considering $\mathbf{Ç}$ as a function $Ç(\mathbf{R})$, $R_{P_*} \leqslant \mathbf{R} \leqslant R_{P^*}$, the first statement of the theorem is immediately obtained from Eq. (10). To prove the second statement expand $(\mathbb{D}^2[\mathbf{R} + Ç(\mathbf{R})s + o\{s\}])^{1/2}$ up to the first power of s. ▮

Proof of Theorem 3.2.1. If Eq. (10) holds, then by renaming $\mathbb{E}[\mathbf{C} \mid R_P]$ as $Ç(P)$ one gets

$$\mathrm{RT}_P(A) = R_P + Ç(P)s + o\{s\},$$

where both R_P and $Ç(P)$ are positive increasing functions of P. Therefore $Ç(P)$ can be considered to be the Pth percentile of a random variable $\mathbf{Ç}$. Its percentiles below P_* and above P^* can be added by an arbitrary continuous extrapolation. Since $\mathbf{R}$ is an increasing function of $\mathbf{Ç}$, one can denote $\inf\{R\}$ by m (a nonnegative value), and present $\mathbf{R}$ as $L(\mathbf{Ç}) + m$, where L is an increasing function vanishing at the lower boundary $\inf\{Ç\}$ of $\mathbf{Ç}$. Renaming $Çs$ as ${}^\circ T(A, Ç)$, one gets

$$\mathrm{RT}_P(A) = {}^\circ T(A, Ç_P) + L(Ç_P) + m + o\{s\},$$

which is asymptotically equivalent to Eq. (8). ▮

Proof of Theorem 3.3.1. The percentiles RT_P and R_P are defined from

$$\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P\}\alpha + \mathrm{Prob}\{\mathbf{H} < \mathrm{RT}_P\}(1-\alpha) = P; \qquad (*1)$$

$$\mathrm{Prob}\{\mathbf{R} < R_P\}\alpha + \mathrm{Prob}\{\mathbf{H} < R_P\}(1-\alpha) = P. \qquad (*2)$$

First, we subject $\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P\}$ to transformations analogous to those used in Theorem 3.1.1:

$$\begin{aligned}
&\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P(A)\} \\
&= \int_0^{\mathrm{RT}_P} \mathbb{F}_R(\mathrm{RT}_P - T)\, d\mathbb{F}_T(T) \\
&= \int_0^{\mathrm{RT}_P} \mathbb{F}_R(\mathrm{RT}_P)\, d\mathbb{F}_T(T) - \int_0^{\mathrm{RT}_P} T\mathbb{f}_R(\mathrm{RT}_P)\, d\mathbb{F}_T(T) + o\{s\}.
\end{aligned}$$

Observing, as in Theorem 3.1.1, that $\mathbb{f}_R(\mathrm{RT}_P)$ can be replaced by $\mathbb{f}_R(R_P)$, and the upper integration limits can be replaced by ∞, the transformations continue as

$$\begin{aligned}
&\mathbb{F}_R(\mathrm{RT}_P) \int_0^\infty d\mathbb{F}_T(T) - \mathbb{f}_R(R_P) \int_0^\infty T\, d\mathbb{F}_T(T) + o\{s\} \\
&\quad = \mathbb{F}_R(\mathrm{RT}_P) - \mathbb{f}_R(R_P)\, \mathbb{E}[\mathbf{T}] + o\{s\}. \qquad (*3)
\end{aligned}$$

Substituting (*3) for $\mathrm{Prob}\{\mathbf{R} + \mathbf{T} < \mathrm{RT}_P\}$ in (*1), we come to the following expression:

$$\mathbb{F}_R(\mathrm{RT}_P)\alpha - \mathbb{f}_R(R_P)\, \mathbb{E}[\mathbf{T}]\alpha + \mathbb{F}_H(\mathrm{RT}_P)(1-\alpha) + o\{s\} = P. \qquad (*4)$$

Due to (*2), we have also

$$\mathbb{F}_R(R_P)\alpha + \mathbb{F}_H(R_P)(1-\alpha) = P. \tag{*5}$$

Subtracting (*5) from (*4), presenting $\mathbb{F}_R(\mathrm{RT}_P) - \mathbb{F}_R(R_P)$ as $\mathbb{f}_R(R_P)(\mathrm{RT}_P - R_P) + \mathrm{o}\{\mathrm{RT}_P - R_P\}$, presenting $\mathbb{F}_H(\mathrm{RT}_P) - \mathbb{F}_H(R_P)$ as $\mathbb{f}_H(R_P)(\mathrm{RT}_P - R_P) + \mathrm{o}(\mathrm{RT}_P - R_P\}$, and observing that $\mathrm{o}\{\mathrm{RT}_P - R_P\}$ is $\mathrm{o}\{s\}$, one gets

$$\{\alpha\mathbb{f}_R(R_P) + (1-\alpha)\,\mathbb{f}_H(R_P)\}(\mathrm{RT}_P - R_P) = \alpha\mathbb{f}_R(R_P)\,\mathbb{E}[\mathbf{T}] + \mathrm{o}\{s\}.$$

Since $\mathbb{E}[\mathbf{T}] = \mathbb{E}[\mathbf{C}]s + \mathrm{o}\{s\}$, the statement of the theorem is obtained from here by algebraic transformations. ∎

Proof of Lemma 3.3.1.

$$\mathbb{E}[\mathbf{RT}] = \{\alpha\mathbb{E}[\mathbf{R}] + (1-\alpha)\,\mathbb{E}[\mathbf{H}]\} + \alpha\mathbf{E}[\mathbf{C}]s + \mathrm{o}\{s\},$$

hence asymptotically the $\mathbb{E}[\mathbf{RT}]$ vs s slope equals $\alpha\mathbb{E}[\mathbf{C}]$. From Eq. (12), the asymptotic RT_P vs s slope equals $\mathbb{E}[\mathbf{C}][1 + \lambda(R_P)(1-\alpha)/\alpha]^{-1}$. This value cannot exceed $\mathbb{E}[\mathbf{C}]$ because $[1 + \lambda(R_P)(1-\alpha)/\alpha]^{-1} \leqslant 1$. Hence the ratio of the two asymptotic slopes cannot be less than $\alpha\mathbb{E}[\mathbf{C}]/\mathbb{E}[\mathbf{C}] = \alpha$. ∎

Proof of Theorem 3.3.2. The theorem states that the race model is asymptotically equivalent to the 2-state mixture model

$$\mathbf{RT}(A) = \begin{cases} \mathbf{N} + \mathbf{R} & \text{with probability } 1-\alpha \\ \mathbf{T} + \mathbf{R} & \text{with probability } \alpha, \end{cases}$$

where α is $\mathrm{Prob}\{\mathbf{N} > 0\}$. $\mathrm{Prob}\{\mathbf{RT} < X\}$ in this model can be presented in the following form:

$$\mathrm{Prob}\{\mathbf{RT} < X\}_{\text{2-state}} = \int_0^\infty \int_{-\infty}^0 \mathbb{F}_R(X-N)\,d\mathbb{F}_{N|T}(N|T)\,d\mathbb{F}_T(T)$$
$$+ \int_0^\infty \int_0^\infty \mathbb{F}_R(X-T)\,d\mathbb{F}_{N|T}(N|T)\,d\mathbb{F}_T(T). \tag{*1}$$

The analogous expression for the race model is

$$\mathrm{Prob}\{\mathbf{RT} < X\}_{\text{race}} = \int_0^\infty \int_{-\infty}^0 \mathbb{F}_R(X-N)\,d\mathbb{F}_{N|T}(N|T)\,d\mathbb{F}_T(T)$$
$$+ \int_0^\infty \int_0^T \mathbb{F}_R(X-N)\,d\mathbb{F}_{N|T}(N|T)\,d\mathbb{F}_T(T)$$
$$+ \int_0^\infty \int_T^\infty \mathbb{F}_R(X-T)\,d\mathbb{F}_{N|T}(N|T)\,d\mathbb{F}_T(T). \tag{*2}$$

The difference between (*2) and (*1) is

$$\int_0^\infty \int_0^T \mathbb{F}_R(X-N)\, d\mathbb{F}_{N\mid T}(N\mid T)\, d\mathbb{F}_T(T) - \int_0^\infty \int_0^T \mathbb{F}_R(X-T)\, d\mathbb{F}_{N\mid T}(N\mid T)\, d\mathbb{F}_T(T). \qquad (*3)$$

Expanding both $\mathbb{F}_R(X-T)$ and $\mathbb{F}_R(X-N)$ up to the first power of T and N, respectively, and observing that $o\{T\} = o\{s\}$ and, since $N < T$ within the integration limits, $o\{N\} = o\{s\}$, the difference (*3) can be rewritten as

$$\mathbb{f}_R(X) \int_0^\infty \int_0^T (T-N)\, d\mathbb{F}_{N\mid T}(N\mid T)\, d\mathbb{F}_T(T) + o\{s\}$$

$$= \mathbb{f}_R(X) \int_0^\infty T \int_0^T (1 - N/T)\, d\mathbb{F}_{N\mid T}(N\mid T)\, d\mathbb{F}_T(T) + o\{s\}.$$

The inner integral is obviously $o\{1\}$, and since T is of the order of s, the entire expression is $o\{s\}$. This proves the statement of the theorem. ▮

Acknowledgments

I am indebted to R. Duncan Luce for many thought-provoking comments and constructive criticism. I am equally indebted to Robert Sekuler for his cooperation in the RT experiments, supported by a grant to Robert Sekuler from the U. S. Air Force Office of Scientific Research, AFOSR 85-0370. I am grateful to James Townsend, Richard Schweickert, Jan P. H. van Santen, Lawrence Hubert, and Joseph Malpeli for many helpful suggestions.

References

Allik, J., & Dzhafarov, E. N. (1984). Reaction time to motion onset: Local dispersion model analysis. *Vision Research*, **24**, 99–101.

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, **21**, 93–123.

Ball, K., & Sekuler, R. (1980). Models of stimulus uncertainty in motion perception. *Psychological Review*, **87**, 435–469.

Chocholle, R. (1940). Variations des temps de réaction auditifs en fonction de l'intensité a diverse frequences. *L'Anne Psychologique*, **41**, 65–124.

Dzhafarov, E. N. (1982). General model for visual motion detection. *Studia Psychologica*, **24**, 193–198.

Dzafarov, E. N., & Allik, J. (1984). A general theory of motion detection. In M. Rauk (Ed.), *Computational Models in Hearing and Vision*, pp. 77–84. Tallin: Estonian Academy of Sciences.

Green D. M., & Luce, R. D. (1971). Detection of auditory signals presented at random times, III. *Perception and Psychophysics*, **9**, 257–268.

Grice, G. R. (1968). Stimulus intensity and response evocation. *Psychological Review*, **75**, 359–373.

Grice, G. R. (1972). Application of a variable criterion model to auditory reaction time as a function of the type of catch trial. *Perception and Psychophysics*, **12**, 103–107.

GRICE, G. R., NULLMEYER, R., & SPIKER, V. A. (1982). Human reaction time: Toward a general theory. *Journal of Experimental Psychology: General*, **111**, 135–153.

KOHFELD, D. L., SANTEE, J. L., & WALLACE, N. D. (1981a). Loudness and reaction time, I. *Perception and Psychophysics*, **29**, 535–549.

KOHFELD, D. L., SANTEE, J. L., & WALLACE, N. D. (1981b). Loudness and reaction time, II. Identification of detection components at different intensities and frequencies. *Perception and Psychophysics*, **29**, 550–562.

LUCE, R. D. (1986). *Response times*. New York: Oxford Univ. Press.

LUCE, R. D. (1991). Personal communication.

LUCE, R. D., & GREEN, D. M. (1972). A neutral timing theory for response times and the psychophysics of intensity. *Psychological Review*, **79**, 14–57.

MCGILL, W. (1963). Stochastic latency mechanisms. In R. D. Luce & E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol. 1, pp. 309–360. New York: Wiley.

NAKAYAMA, K. (1985). Biological image motion processing: A review. *Vision Research*, **25**, 625–660.

PIERON, H. (1920). Nouvelles recherches sur l'analyse du temps de latence sensorielle et sur la loi qui relie le temps à l'intensité d'excitation. *L'Anne Psychologique*, **22**, 58–142.

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

REICHARDT, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communications*, pp. 303–317. Cambridge, Massachusetts: MIT Press.

TOWNSEND, J. T., & ASHBY, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cabridge Univ. Press.

TROSCIANKO, T., & FAHLE, M. (1988). Why do isoluminant stimuli appear slower? *Journal of the Optical Society of America*, **A5**, 871–879.

RECEIVED: May 16, 1990