Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes

Haiyuan Yang [a,*], Mario Fific [b], James T. Townsend [a]

[a] *Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, 47405, USA*
[b] *Psychology Department, Grand Valley State University, MI, 49401, USA*

HIGHLIGHTS

- We explore the precise behavior of the serial exhaustive SIC function for $n = 2$.
- We provide a generalization of the SIC function to an arbitrary number of processes.
- We analyze the generalized SIC for both parallel and serial models with minimum and maximum time stopping rules.
- We demonstrate application of the theorems to data from a short-term memory search task.

ARTICLE INFO

ABSTRACT

The Survivor Interaction Contrast (SIC) is a distribution-free measure for assessing the fundamental properties of human information processing such as architecture (i.e., serial or parallel) and stopping rule (i.e., minimum time or maximum time). Despite its demonstrated utility, there are some vital gaps in our knowledge: first, the shape of the serial maximum time SIC is theoretically unclear, although the one 0-crossing negative-to-positive signature has been found repeatedly in the simulations. Second, the theories of SIC have been restricted to two-process cases, which restrict the applications to a limited class of models and data sets. In this paper, we first prove that in the two-process case, a mild condition known as strictly log-concavity is sufficient as a guarantor of a single 0-crossing of the serial maximum time SIC. We then extend the definition of SIC to an arbitrary number of processes, and develop implicated methodology of SIC in its generalized form, again in a distribution-free manner, for both parallel and serial models in conjunction with both the minimum time and maximum time stopping rules. We conclude the paper by demonstrating application of the theorems to data from a short-term memory search task.

Published by Elsevier Inc.

1. Introduction

The question of whether people can perform multiple perceptual or mental operations simultaneously, that is, parallel processing, vs. whether items or tasks must proceed serially (one at a time), has intrigued psychologists since the birth of experimental psychology. Historically, reaction time (RT) has been the primary measure on this question. The work of the physiologist F.C. Donders (e.g., Donders, 1868) was seminal in this regard, although other researchers, such as W. Wundt, were more prolific with regard to early results on human cognition.

With the revolution brought about through cognitive science and cognitive psychology in the 1950s and 1960s, questions such as the parallel vs. serial conundrum, which had lain dormant since the nineteenth century saw a renaissance of interest.

The serial vs. parallel topic is our primary concern here. However, it may be worth a moment's pondering, given the pioneering role of William K. Estes in the advent of mathematical psychology, of how the latter field, and Estes' research, fit into, and contributed to, modern cognitive psychology. Three tributaries fed the new stream of mathematical psychology in the 1950s and 60s. These were: 1. Signal detection theory, child of psychophysics and sensory processes, mathematical communications theory, applied physics, and statistical decision making (e.g., Green & Swets, 1966; Tanner & Swets, 1954). 2. Foundational measurement the offspring of S.S. Stevens' brilliant but non-rigorous statements concerning measurement in psychology fostered and rendered rigorous through strands from philosophy, mathematical logic and abstract algebra (e.g., Krantz, Luce, Suppes, & Tversky, 1971; Roberts & Zinnes, 1963). 3. Mathematical learning theory which went back at least to Clark Hull (e.g., Hull, 1952); or see Koch's elegant summary in Modern Learning Theory (Koch, 1954). This branch is where we find the Estes trailblazing Stimulus Sampling Theory (Estes, 1955, 1959), a precise, quantitative theory of human and animal learning.

---

* Corresponding author.
E-mail addresses: haiyyang@indiana.edu (H. Yang), fificm@gvsu.edu (M. Fific), jtownsen@indiana.edu (J.T. Townsend).

This theory, which still impacts a wide spectrum of research in cognition today, led to a score of research advances by Estes and colleagues as well as a host of other scientists (e.g., Atkinson & Estes, 1963; Friedman et al., 1964).

Estes was an early entrant into the embryonic cognitive movement. His research in this domain was likely influenced by the burgeoning efforts utilizing the information processing approach, perhaps the early dominant theme in this new domain. Early pioneers included Wendell Garner (e.g., Garner, 1962), Donald Broadbent (e.g., Broadbent, 1958), William Hick (e.g. Hick, 1952), and Colin Cherry (e.g. Cherry, 1953) (note the heavy presence of British psychologists).

American psychologists were soon contributing to this rapidly expanding field which bridged sensory processes, higher perception, and elementary cognition. Prime examples are Charles Eriksen (e.g., Eriksen & Spencer, 1969), Michael Posner (e.g., Posner, 1978), Raymond Nickerson (e.g., Nickerson, 1972), Ralph Haber (e.g., Haber & Hershenson, 1973), and Howard Egeth (e.g., Egeth, 1966). And, Bill Estes of course.

The employment of ingenious experimental designs to answer questions concerning whether humans perform visual or memory search in a serial or parallel fashion provide apt examples of new trends making an appearance in the 60s and 70s. (e.g., Sperling, 1960, 1967; Sternberg, 1966, 1975). Estes and colleagues provided some classic early results in this domain in extending, and mathematically modeling extensions of Sperlings innovative visual search experimental designs. For instance, Estes and Taylor (1964) developed a new detection method as well as associated models in this vein. Also, Estes and Taylor (1966) and Estes and Wessel (1966) were beginning to explore phenomena and human information processing mechanisms related to the presence of redundant signals in visual displays.

The Sternberg (1966) innovative and rather startling RT data in short term memory search, in particular, had a profound influence on thinking in the parallel vs. serial processing literature. In fact, a massive body of experimental literature over several decades has been based on the inference that increasing, more-or-less straight-line RT functions of the workload n,[1] the number of comparisons to perform, imply serial processing. However, the ability of limited capacity parallel models to mimic serial models, in the strong sense of mathematical equivalence, was demonstrated relatively early on (e.g., Atkinson, Holmgren, & Juola, 1969; Murdock, 1971; Townsend, 1969, 1971).[2] And in fact, the reverse possibility of serial models to mimic parallel models was also proven (Townsend, 1969, 1971, 1972, 1974). The early mathematical results were confined to limited types of RT distributions, but later developments extended to arbitrary probability distributions (Townsend, 1976; Townsend & Ashby, 1983; Vorberg, 1977).

The parallel models which perfectly mimic serial models are *limited capacity* in the sense that their processes degrade in their efficiency as the workload n increases. Such models intuitively make the predictions associated with serial processing, specifically the linear RT graphs of the workload n (e.g., Townsend, 1971). Fortunately, theory-driven experimental methodologies have been invented in recent years that are considerably more robust in the

assessment of mental architecture, particularly serial vs. parallel processing (Scharff, Palmer, & Moore, 2011; Townsend, 1976, 1981, 1990a; Townsend & Nozawa, 1995; Townsend & Wenger, 2004). In particular, the new methodologies often allow architectural inferences even though the workload is held constant, so that capacity does not confound architectural inferences.

Our focus here lies within the general approach referred to as *Systems Factorial Technology* (hereafter SFT; see Townsend, 1992; Townsend & Nozawa, 1995). A number of investigators have made essential contributions to this literature including Schweickert and Dzhafarov and colleagues (Dzhafarov, 1997; Dzhafarov, Schweickert, & Sung, 2004; Schweickert, 1978, 1982; Schweickert & Giorgini, 1999; Schweickert, Giorgini, & Dzhafarov, 2000). SFT relies heavily on mathematical propositions indicating experimental conditions where strong tests of architectures may be found, although other testable features, such as capacity, are also encompassed presently. The bulk of theoretical work has been performed under the assumption of *selective influence*. Our scope prohibits details here, but we can loosely define selective influence as the property that certain experimental factors act only on specific processes in the overall system (see Section 2.1 for more detailed discussion on selective influence). When selective influence is in force, predictions of serial and parallel models and the pertinent decisional stopping rules are strikingly distinct. This paper is intended to significantly strengthen and extend these predictions.

SFT requires the survivor function $S(t)$, which is simply the complement of the well-known cumulative distribution (or frequency) function (the CDF) written as $F(t)$. That is, $S(t) = 1 - F(t)$. A central statistical diagnostic is then the survivor interaction contrast (or SIC) function. It performs a double difference contrast operation on the survivor functions that is analogous to the mean interaction contrast (or MIC) employed on the arithmetic RT means in earlier investigations (e.g., Schweickert, 1978; Sternberg, 1966). However, it now expresses a highly diagnostic function of time, rather than a single number.

Despite the successful deployment of the SIC measure, there are some vital gaps in our knowledge, restricting the applications to a limited class of models and data sets. These will be sketched within a brief presentation of relevant knowledge we do have.

We know that, for $n = 2$, serial minimum time models predict perfectly flat signatures whereas serial maximum time (i.e., the classical exhaustive processing time stopping rule; see Sternberg, 1969; Townsend, 1974) predictions must include at least one wiggle (i.e., the up-and-down excursions marked by 0-crossings) below and above 0 (Townsend & Nozawa, 1995). However, although simulations have intimated that there is a single wiggle passing through 0, in a negative-to-positive direction as exhibited in Fig. 1 (the top right panel), this has not been shown to be true for all distributions. In fact, the exact shape of the SIC curve is as yet unknown.

Therefore, in elucidating further properties of serial exhaustive processing: A. We first prove that serial exhaustive processing inevitably predicts an odd number of 0-crossings in the $n = 2$ case. B. Next, we show that a certain readily-met mathematical condition is sufficient to force the behavior indicated through our simulations, a single 0-crossing of the SIC function.

The behavior our SIC signatures have also remained unidentified for $n > 2$, in the case of all studied serial and parallel processes up to now. The quite intriguing behaviors in the case of the serial and parallel models with varying stopping rules, and for arbitrary values of n, are next developed for: A. Serial minimum time processing. B. Serial maximum time processing. C. Parallel minimum time processing. D. Parallel maximum time processing.

Successful completion of the above goals should significantly expand the possibilities of application. Since mathematical details of SFT in general, have been published elsewhere (e.g., Townsend & Nozawa, 1995) and tutorials are available (e.g., Townsend, Fific, & Neufeld, 2007; Townsend & Wenger, 2004; Townsend et al., 2011), only the bare bones SFT can be displayed here.

---

[1] An increment in workload is usually natural to define in terms of number of dimensions, or subtasks involved in some task. We shall often refer simply to items or, sometimes, processes as generic tags for the discrete objects being processed or the conduits working on them. The unit of workload typically relates in a natural fashion to the task. For example, if a memory search task involves examination of a list of letters, the unit may be made straightforwardly in terms of letters. Then n may stand for both the workload in the task and the number of letters in the memory set.

[2] For an up to date review of the parallel–serial identifiability issue, see Townsend, Yang, and Burns (2011).
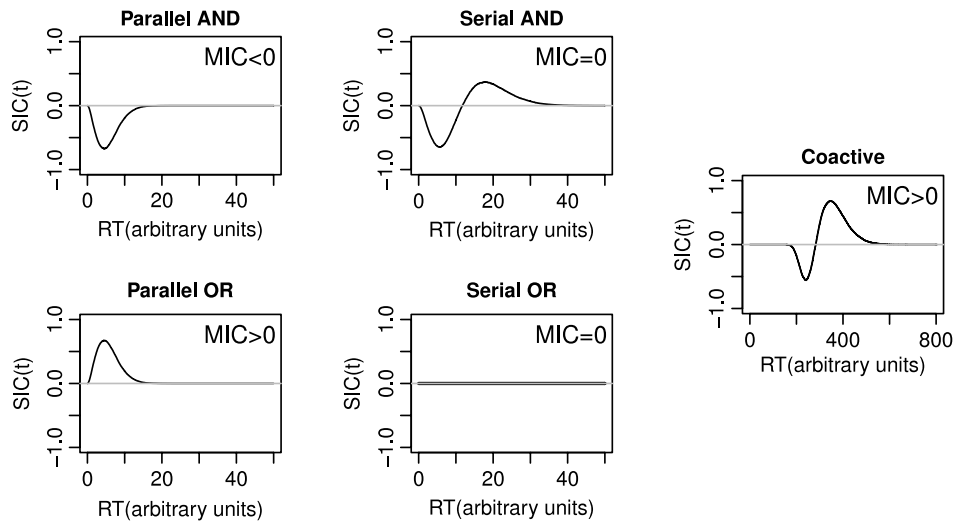
Fig. 1. Representative predictions of survivor interaction contrast (SIC) and mean interaction contrast (MIC) across different architectures and stopping rules. The four canonical models on the left are based on simulations of gamma distributions, with various shape parameters regarding different salience levels of processes. The coactive model on the right is based on simulations of dynamic systems described in (Townsend & Wenger, 2004).
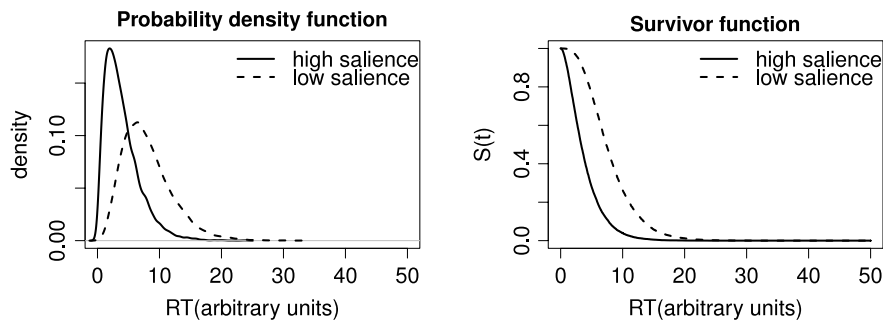

Fig. 2. Distributional influences of the high vs. low salience manipulation with survivor functions (right) and densities suggesting satisfaction of the one point crossover (left). The simulations are based on gamma distributions with different shape parameters.

2. Extending knowledge of serial exhaustive processing: analysis for two processes

2.1. Underlying assumptions of SFT

The critical assumption underlying SFT is selective influence, which means that an experimental factor affects only a single process. We continued the convention (Townsend & Ashby, 1983) of indicating the level of a factor as high if it speeds up the process and as low when it slows down the process. Manipulating the speed of the process is also referred to as salience manipulation. In the short-term memory search task shown in the exemplar experiment section, for instance, high dissimilarity between the item in the memory and the target will be designated as a high salience level, and low dissimilarity between item and target will be designated as a low salience level.

In terms of RT distribution, an experimental variable acting selectively could in principle affect any one or more of several aspects of an RT distribution. In this paper we follow the assumption as expressed in Townsend and Nozawa (1995) that the two processing time density functions for a single process, at the two factor levels cross exactly once, i.e., there is exactly one time t^*, where the two density functions are equal. This assumption implies an ordering of both survivor functions (i.e. stochastic dominance) as well as the means and medians. This assumption appears to be satisfied within the limits of typical psychological applications (e.g., Townsend, 1990b; Townsend & Nozawa, 1995). Fig. 2 shows a hypothetical example of selective influence operating at the one-point density

crossover level, using two gamma distributed variables. We can see that the property of single density crossover (left panel) leads to both an ordering of the survivor functions (right panel) and of the mean processing times (right panel, the area under the survivor curves).

In order to avoid possible failure of selective influence due to indirect non-selective influence, i.e., one factor can indirectly affect the 'wrong' process through stochastic dependence (Townsend, 1984; Townsend & Thomas, 1994), some condition concerning stochastic independence is typically required. For simplicity of proof, stochastic independence of the process completion times is assumed, although Dzhafarov (1999, 2003) and Dzhafarov et al. (2004) has shown that conditional independence given some common sources of randomness which are unaffected by the experimental factors is sufficient for these purposes.[3]

Therefore, not only must selective influence act at a level of sufficient power on a process X (e.g., the density single point crossing assumption above), but it must be assumed that the marginal probability functions on processing times for all other processes Y, Z, etc. must be invariant when the X factor is manipulated. Base times (all else besides the completion times of the processes in which we are interested) are avoided in this study but under the assumption of conditional independence, would not affect our results in any event.

---

[3] See Kujala and Dzhafarov (2008) for recent advances pertaining to selective influence.

2.2. Limited review of systems factorial technology

Suppose we are concerned with just two processes or channels. We refer to these as X and Y. Let $f_{XL}(t)$ be the density function of the processing time on X when the factor level of process X is low and $f_{XH}(t)$ be the density function of the processing time on X when the factor level is high. Likewise, let $f_{YL}(t)$ and $f_{YH}(t)$ be the density functions of the processing time on Y when the factor level of Y is low and high, respectively. In this paper we assume that all density functions are sufficiently smooth.[4] Let $S_{XL}(t)$, $S_{XH}(t)$, $S_{YL}(t)$ and $S_{YH}(t)$ denote the survivor functions corresponding to $f_{XL}(t)$, $f_{XH}(t)$, $f_{YL}(t)$ and $f_{YH}(t)$ respectively. A survivor function of random variable T is defined as

$$S(t) = P(T > t) = \int_t^\infty f(t')dt' = 1 - F(t). \quad (1)$$

The Survivor Interaction Contrast (SIC) function of the total reaction time T (later we use T with subscripts to denote reaction times of single process) is defined as

$$\text{SIC}^2(t) = (S_{LL}(t) - S_{LH}(t)) - (S_{HL}(t) - S_{HH}(t)). \quad (2)$$

The superscript indicates the number of processes, and subscripts are used to denote the salience level of each process. For example, $S_{LL}(t)$ indicates the survivor function of RT for the condition in which both process X and Y are of low salience. For simplicity, we use Δ^2 to denote the double differences over the factor level, hence Eq. (2) becomes $\text{SIC}^2(t) = \Delta^2_{X,Y}S(t)$. Theoretically, the observed RT should be a combination of process completion times, with appropriate forms (i.e. sum, max, min, or probability mixture) depending on the underlying stopping rules.

The mean interaction contrast (MIC) is also an important statistic in distinguishing among certain processing types. With $\overline{RT}$ indicating the mean response time and the subscripts as defined above, the MIC is given by

$$\text{MIC} = (\overline{RT}_{LL} - \overline{RT}_{LH}) - (\overline{RT}_{HL} - \overline{RT}_{HH}). \quad (3)$$

Sternberg (1969) suggested that based on selective influence, serial models with independent processing times would exhibit MIC = 0. The use of MIC has later been extended to diagnose parallel processing (Schweickert & Townsend, 1989; Townsend & Nozawa, 1995). The SIC and MIC predictions[5] of the four standard models are shown in Fig. 1. Note that, due to the fact that the integral of the survivor function of a positive random variable is its expected value, the interaction contrast of mean values is equal to the integrated SIC, i.e., $\int_0^\infty \text{SIC}(t)dt = \text{MIC}$.

Fig. 1 shows that parallel-processing SICs reveal total positivity (i.e. the SICs are non-negative functions of t) in the case of minimum time conditions (MIC > 0) but total negativity in the case of maximum time conditions (MI < 0). On the other hand, serial minimum time SICs are equal to zero at all time values t (MIC = 0), while simulation results have repeatedly found, in contrast, that serial maximum time (or serial exhaustive) SICs show a large negative portion, followed by an equally large positive portion (MIC = 0). The coactive model, based on the summed Poisson processes, predicts that the SIC is negative for small times and then positive for later times, much like the SIC for the serial exhaustive model. But the negative region is always smaller than the positive region, which leads to a positive MIC value[6] (MIC > 0). Thus each of the five models makes a unique prediction for the combination of MIC value and SIC shape. By using both the MIC and SIC statistics, one can differentiate between serial, parallel, and coactive architectures, as well as minimum time and exhaustive stopping rules.

As observed earlier, for the case of serial exhaustive processing, Townsend and Nozawa (1995) proved that the SIC function begins negative but must be positive for substantial values of $t > 0$. In fact, the summation of the positive portion of SICs is equal to the summation of the negative portion, since the mean interaction contrast must be zero. However, it does not follow from the existing proofs that the SIC must be the 1-wiggle S-shape function as found in the simulations of Fig. 1.

Can it be demonstrated that the 1-wiggle SIC behavior applies to all distributions when convolved to produce serial exhaustive predictions? This might seem rather unlikely given that, in principle the underlying distributions are arbitrary. Are there nontrivial conditions that are necessary and/or sufficient to elicit the 1-wiggle portrait? In attempting to answer these questions, we have first of all discovered novel aspects of general wiggle behavior not only for the involved processes of $n = 2$ but for arbitrary $n \geq 2$.

We are now prepared for our first theoretical result.

2.3. Theoretical propositions

Although we know from previous work (Townsend & Nozawa, 1995) that wiggles must exist for 2-stage serial exhaustive processing, we do not know how many in general we should expect. Then our first result demonstrates that for any underlying processing distributions in series, the number of crossovers of 2-stage serial exhaustive SIC must be an odd number.[7]

Proposition 2.1. *Assume selective influence. The independent serial exhaustive SIC must have an odd number of crossovers with horizontal axis in the interval $(0, +\infty)$.*

Proof. The overall RT in serial exhaustive models should be the summation of the completion times of two processes X and Y, i.e., $T = T_X + T_Y$. Since we assume that two channels process independently, we can write the survivor interaction contrast as

$$\text{SIC}^2_{\text{ser.AND}}(t) = \Delta^2_{X,Y}P(T_X + T_Y > t)$$

$$= -\Delta^2_{X,Y}\int_0^t F_X(t - t_y) \times f_Y(t_y)dt_y$$

$$= -\int_0^t \Big[F_{XL}(t - t_y) \times f_{YL}(t_y) - F_{XL}(t - t_y) \times f_{YH}(t_y)$$
$$- F_{XH}(t - t_y) \times f_{YL}(t_y) + F_{XH}(t - t_y) \times f_{YH}(t_y) \Big]dt_y$$

$$= -\int_0^t \Big[[F_{XH}(t - t_y) - F_{XL}(t - t_y)]$$
$$\times [f_{YH}(t_y) - f_{YL}(t_y)] \Big]dt_y. \quad (4)$$

Because of the action of selective influence on the survivor functions, the first multiplicand under the integral sign is non-negative. Further, when selective influence operates at the one-point density crossover level, the second multiplicand will be positive for $t < t^*$, where t^* represents the density crossover point. Thus SIC function must be negative for small times $t < t^*$.

To show that there are odd number crossovers, we need to demonstrate that when $t \to \infty$, the $\text{SIC}^2(t)$ must converge to zero

---

[4] A smooth function of class C^k is a function that has continuous derivatives up to the kth order in its domain. Here the term "sufficiently smooth" means that the density functions do not need to be of class C^∞, but all operations on the density functions mentioned in this paper should be well defined.

[5] For rigorous proofs of SIC predictions, see Townsend and Nozawa (1995); Townsend and Wenger (2004).

[6] Although both coactive models and parallel race models predict everywhere-positive OR MIC results, an investigator might test them using a result of

Schweickert and Wang (1993). Namely, if the factor levels are increased over a sufficiently large range, parallel race models predictions will approach a MIC limit whereas coactive models predictions will not.

[7] In this paper, we restricted our discussion to finite-zero cases, i.e., the number of crossovers of SIC would be either even or odd.

from the positive side. Let $DF_X(t) = F_{XH}(t) - F_{XL}(t)$, $Df_Y(t) = f_{YH}(t) - f_{YL}(t)$. Because of the property of commutativity of convolution, we can rewrite $SIC^2(t)$ as

$$SIC^2_{ser.AND}(t) = -\int_0^t DF_X(t - t_y) \times Df_Y(t_y) dt_y$$

$$= -\int_0^t DF_X(t_y) \times Df_Y(t - t_y) dt_y. \quad (5)$$

Denote $I^2(s) = \int_0^s SIC^2_{ser.AND}(t) dt$. By Fubini's theorem,

$$I^2(s) = -\int_0^s \int_0^t DF_X(t_y) \times Df_Y(t - t_y) dt_y dt$$

(Change the order of integration)

$$= -\int_0^s \int_{t_y}^s DF_X(t_y) \times Df_Y(t - t_y) dt dt_y$$

(Definition of density function)

$$= -\int_0^s DF_X(t_y) \times DF_Y(s - t_y) dt_y. \quad (6)$$

Because of the action of selective influence on the survivor functions, both $DF_X(t)$ and $DF_Y(t)$ are positive, thus $I^2(s)$ is negative. Another fact about $I^2(s)$ is that it goes to zero when $s \to \infty$ (recall that $I^2(\infty) = $ MIC). Combining these two facts together we see that $I^2(s)$ must increase toward zero from the negative side at the end. Thus the derivative of $I^2(s)$, i.e. $SIC^2(s)$, must be positive as long as s is large enough.

Having shown that the SIC function is negative for small time values and positive when it approaches to zero at the end, the implication is that the SIC must have an odd number of crossovers in $(0, +\infty)$. □

Many simulations with differing distributions have intimated that perhaps the odd number of crossovers might typically just be 1. It is next proven that a satisfyingly weak condition ensures that this will be the case. However, first a brief discussion of the relevant history is in order.

From Eq. (5) we see that the serial exhaustive SIC curve is a convolution of the difference in probability density functions, of the high salience minus the low salience on the X factor, and the difference in cumulative distribution functions of the high salience minus the low salience on the Y factor. Further, the integral of the $SIC^2(t)$ with variable upper limit, i.e., $I^2(s)$, is also a convolution, of two functions which are the two differences in cumulative distribution functions on X and Y factor respectively. For many years, it was thought that a condition of unimodality[8] of such functions was enough to produce our needed result of a single 0-crossing. This was eventually found to be false. That is, it can be shown that unimodality of both functions in the convolution is not sufficient to imply unimodality of the integral (Chung, 1953). In contrast, Ibragimov (1956) proposed and demonstrated that the convolution of any two unimodal functions will be still unimodal, if at least one of them is logarithmic concave (so called strong unimodal).[9]

Proposition 2.2. *Assume selective influence. The independent serial exhaustive SIC crosses the time axis only once in the interval $(0, +\infty)$, if either $F_{XH}(t) - F_{XL}(t)$ or $F_{YH}(t) - F_{YL}(t)$ is strictly log-concave.*

Proof. Let $DF_X(t) = F_{XH}(t) - F_{XL}(t)$, $DF_Y(t) = F_{YH}(t) - F_{YL}(t)$. Also let $Df_X(t) = f_{XH}(t) - f_{XL}(t)$, $Df_Y(t) = f_{YH}(t) - f_{YL}(t)$. Assume that the two processing time density functions at the two factor levels ($f_{XL}(t)$ and $f_{XH}(t)$; $f_{YL}(t)$ and $f_{YH}(t)$) cross exactly once, thus both $DF_X(t)$ and $DF_Y(t)$ are strictly unimodal. Suppose $DF_Y(t)$ is strictly log-concave. Assume also, for the moment, that $DF_Y(t)$ has support $(-\infty, \infty)$, i.e., $DF_Y(t) \neq 0$ for all $t \in (-\infty, \infty)$ (we will remove this assumption later). We show that the $SIC^2(t)$ has only one 0-crossing for any $DF_X(t)$. This proof is based on Chapter 1 of Dharmadhikari and Joag-Dev (1988).

Since convolution commutes with translations, we assume, without loss of generality, that $DF_X(t)$ is unimodal about 0, i.e., $Df_X(t) > 0$ for $t < 0$ and $Df_X(t) < 0$ for $t > 0$. Recall from Eq. (5) that $SIC^2_{ser.AND}(t) = -\int_{-\infty}^\infty Df_X(t_y) \times DF_Y(t - t_y) dt_y$. Then for any t, z such that $t < z$, we have

$$SIC^2_{ser.AND}(z) = -\int_{-\infty}^\infty Df_X(t_y) \frac{DF_Y(z - t_y)}{DF_Y(t - t_y)} \times DF_Y(t - t_y) dt_y. \quad (7)$$

Since $DF_Y(t)$ is strictly log-concave, $DF_Y(t + \delta)/DF_Y(t)$ is decreasing in t if $\delta > 0$ and increasing in t if $\delta < 0$. Therefore,

$$\frac{DF_Y(z - t_y)}{DF_Y(t - t_y)} < \frac{DF_Y(z)}{DF_Y(t)} \quad \text{if } t_y < 0$$

and

$$\frac{DF_Y(z - t_y)}{DF_Y(t - t_y)} > \frac{DF_Y(z)}{DF_Y(t)} \quad \text{if } t_y > 0. \quad (8)$$

But we also know that $Df_X(t) > 0$ for $t < 0$ and $Df_X(t) < 0$ for $t > 0$. Consequently Eq. (7) shows that

$$SIC^2_{ser.AND}(z) > -\frac{DF_Y(z)}{DF_Y(t)} \int_{-\infty}^\infty Df_X(t_y) DF_Y(t - t_y) dt_y$$

$$= \frac{DF_Y(z)}{DF_Y(t)} \times SIC^2_{ser.AND}(t). \quad (9)$$

Thus $SIC^2_{ser.AND}(t) = 0 \Rightarrow SIC^2_{ser.AND}(z) > 0$ for all $z > t$. Therefore, SIC has only one 0-crossing and correspondingly, $I^2(s)$ is strictly unimodal.

The condition that $DF_Y(t)$ is never zero can be removed by constructing a series of strictly log-concave functions $DF_Y^{(m)}(t)$ that all have support $(-\infty, \infty)$ and converge to $DF_Y(t)$.[10] Thus $I^{2(m)}(s) = -\int_0^s DF_X(t_y) \times DF_Y^{(m)}(s - t_y) dt_y$ is strictly unimodal by the above proof. Since the class of strictly unimodal distributions on $\mathbb{R}$ is closed under weak limits (Dharmadhikari & Joag-Dev, 1988, page 3), let $m \to \infty$, $I^2(s)$ is strictly unimodal,[11] and $SIC^2(s)$ has only one 0-crossing. □

We should notice that Proposition 2.2 provides a sufficient but not necessary condition for the single 0-crossing property of the SIC function. Hence even when both $DF_X(t)$ and $DF_Y(t)$ are not strictly log-concave, it is possible that the SIC function still has only one 0-crossing point. Does this result mean that the "log-concavity" condition is too far from necessary regarding this problem? The answer is no. As indicated in Ibragimov (1956), for any unimodal function that is not log-concave, there exists another unimodal function such that the convolution of the two functions is multi-modal. In other words, for any $DF_X(t)$ that is not log-concave, we can find a $DF_Y(t)$ such that the corresponding SIC function has more than one zero.

---

[8] Different sources have slightly different definitions for a unimodal function. We shall use the following: a mode of a function f is a number a such that (a). f is non-decreasing on $(-\infty, a]$ and (b). f is non-increasing on $[a, \infty)$. f is unimodal if it has a mode. f is strictly unimodal if it has a single mode.

[9] "Concavity" of an increasing curve means that it bends downward as it ascends, implying a so-called "negative second derivative" in elementary calculus. "Log-concavity" then simply means that the logarithm of the function rather than the function itself is concave.

[10] For readers who are interested in the details, we recommemd (Dharmadhikari & Joag-Dev, 1988), page 21, in which they show how to approximate the log-concave function g supported in $[a, b]$ by the sequence $g^{(m)}$ supported in $(-\infty, \infty)$.

[11] Although (Dharmadhikari & Joag-Dev, 1988) demonstrated their theorem with distributions, we can easily apply their result to our functions by rescaling $I^{2(m)}(s)$.
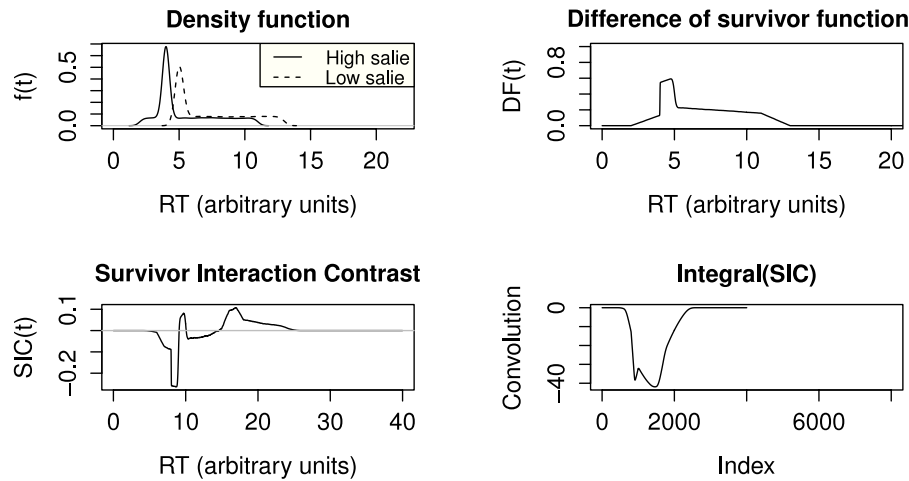
Fig. 3. An example of 2-stage serial exhaustive SIC function that has three 0-crossing points. Graphs include the density functions (top left), the difference of the two corresponding survivor functions (top right), the SIC function (bottom left), and the integral of the SIC function with variable upper limits (bottom right).

We demonstrated the existence of multi-zero SICs by simulations. Completion times on processes were generated for both low and high salience regarding the following rules: 1. For simplicity, we let completion times on processes X and Y have the same distribution form, i.e., $F_{XL}(t) = F_{YL}(t)$, $F_{XH}(t) = F_{YH}(t)$. Thus $DF_X(t) = DF_Y(t)$. 2. The completion times under the two salience conditions are both mixtures of uniform and Gaussian distributions with different parameters designed so that the two density functions at the two factor levels cross exactly once and the corresponding $DF(t)$ is unimodal (but not log-concave). The simulation results are shown in Fig. 3. The top panel shows the two density functions for the low and high conditions on the left and the $DF(t)$ on the right. The bottom left panel shows the SIC function calculated by taking the double differences of survivor functions from the top left panel, and the bottom right panel shows the result of the convolution of $-DF_X(t)$ and $DF_Y(t)$, which is also the integral of the SIC function, as we indicated in the proof above. As we predicted, the convolution function in the bottom right has three (rather than one) turning points, and corresponding with this is that the SIC function on the left has three 0-crossing points. Also note that in the limit, the convolution function in the bottom right equals 0 as it should, indicating the MIC is still 0 as it should be due to seriality. This simulation supports our finding that when losing the property of log-concavity, the serial exhaustive SIC function may exhibit three non-trivial zeros.

It was intimated in the previous section that the condition of log-concavity is not very restrictive. This claim is supported by considering a simple power function (such as assumed in S.S. Steven's famous law): $y = Ax^B$, where A and B are positive constants, x is the independent, and y the dependent variable. For $B \geq 2$, this function is positively accelerated, that is, is convex with the acceleration increasing with B. Yet, however large B may be, it is always log-concave, since $\log(y) = \log(A) + B \times \log(x)$ which is always concave. And as observed, many commonly used distributions obey this caveat, such as the normal distribution, the exponential distribution, and the gamma distribution with shape parameter $P \geq 1$, etc. Proposition 2.2 therefore strongly suggests that we should not be astonished to discover a single 0-crossing when processing is serial with an exhaustive stopping rule.

Nonetheless, as mentioned earlier, the SIC has so far been restricted to the $n = 2$ case. This fact is of more than simple technical interest, since it has excluded the methodology from the larger numbers of items or processes which are often used in perceptual and cognitive experiments. The next part of this paper is devoted to generalizing the knowledge base of the SIC signatures to arbitrary values of n.

3. Fundamental architectural signatures for an arbitrary number of processes

How the two basic architectures, parallel and serial with varying stopping rules behave for $n > 2$ processes will be explored in this section. Before digging into the detail of the extended theory, we first introduce a measure, which is analogous to the $SIC^2(t)$ in Eq. (2), but in a more general form. It will be evident that the complete factorial design can be erected by a recursive embedding in higher and higher values of n, as indicated when $n = 3$. The SIC function of total reaction time T in 3-process case produces the expression

$$SIC^3(t) = [(S_{LLL}(t) - S_{LLH}(t)) - (S_{LHL}(t) - S_{LHH}(t))]$$
$$- [(S_{HLL}(t) - S_{HLH}(t)) - (S_{HHL}(t) - S_{HHH}(t))] \quad (10)$$

where superscript indicates the number of processes and subscripts are used to denote the salience level of each process. For example, $S_{LLL}(t)$ indicates the survivor function of RT for the condition in which all three processes are of low salience. Note that the factorial combination of three factors with their two salience levels leads to eight experimental conditions, thus the SIC function in 3-process case is composed of eight items. An abstract form which is equivalent to Eq. (10) is as follows: $SIC^3(t) = \Delta^3_{X_1, X_2, X_3} S(t)$, where X_i represents the process i. This form of SIC function can be straightforwardly generalized to the case for arbitrary n processes: let us denote the individual process as X_i, $i = 1, 2, \ldots, n$, thus the SIC function in the n-process case could be written as

$$SIC^n(t) = \Delta^n_{X_1, \ldots, X_n} S(t) \quad (11)$$

in which $\Delta^n_{X_1, \ldots, X_n}$ represents the n-order mixed partial difference over the factor levels. Analogous to the 3-process SIC which contains 2^3 items, the n-process SIC function is composed of 2^n survivor functions.

Now we are ready to present the theoretical results for $n > 2$. Our goal is to demonstrate that for the two major stopping rules, minimum and maximum processing times, the results for arbitrary n are close relatives to those for $n = 2$ in the case for both parallel and serial systems.[12] Here we employ mathematical induction in the proofs. Mathematical induction is usually used to establish that a given statement is true for all non-negative integers. It could be done by proving that the first statement in the infinite sequence

---
[12] We limit our scope on pure serial and parallel systems. Mixed system of serial and parallel is not addressed in this paper.

of statements is true, and then proving that if the first N (any arbitrarily chosen number) statements in the infinite sequence of the statements are true, then so is the next one.

We begin with the basic parallel horse race—independent parallel processing and a minimum time stopping rule.

3.1. Parallel minimum time processing

In the case for $n = 2$, the SIC is always positive. Proposition 3.1 reveals that this is the canonical signature for all n.

Proposition 3.1. *Assume selective influence. The independent parallel minimum time processing predicts that the SIC curve will always be positive as a function of time t, for every n.*

Proof. Recall that without the consideration of base time, the total reaction time in parallel minimum time process models should be the minimum of the reaction time on each process i.e, $T = \min(T_{X1}, \ldots, T_{Xn})$. Since we assume that two channels process independently, we have

$$
\begin{aligned}
\mathrm{SIC}^n_{\mathrm{par.OR}}(t) &= \Delta^n_{X_1,\ldots,X_n} P(\min(T_{X1}, \ldots, T_{Xn}) > t) \\
&= \Delta^n_{X_1,\ldots,X_n}[P(T_{X1} > t) \times \cdots \times P(T_{Xn} > t)] \\
&\quad \text{(factoring)} \\
&= [P(T_{XnL} > t) - P(T_{XnH} > t)] \\
&\quad \times \Delta^{n-1}_{X_1,\ldots,X_{n-1}}[P(T_{X1} > t) \times \cdots \times P(T_{X(n-1)} > t)] \\
&= [F_{nH}(t) - F_{nL}(t)] \times \mathrm{SIC}^{n-1}_{\mathrm{par.OR}}(t) \quad (12)
\end{aligned}
$$

where subscripts are used to denote the salience level of each process. For example, T_{XnL} indicates the RT for the condition in which the nth process is of low salience, F_{nL} indicates the corresponding marginal CDF. Because of the action of selective influence on the survivor functions, $F_{nH}(t) > F_{nL}(t)$ for all t. Since $\mathrm{SIC}^2_{\mathrm{par.OR}}(t)$ is always positive (Townsend & Nozawa, 1995), we can infer that $\mathrm{SIC}^n_{\mathrm{par.OR}}(t)$ is always positive, by simple mathematical induction. □

It is interesting that the sign of the parallel horse race SIC function is always positive as n is varied. We will learn that this does not inevitably occur in a parallel system when considering different stopping rules. The next proposition treats maximum time parallel processing.

3.2. Parallel maximum time processing

When $n = 2$, the SIC function is always negative. From the above proposition with minimum time processing we might expect an analogous invariance with a maximum time (i.e., exhaustive processing) stopping rule. Intriguingly, this expectation is not fulfilled as the following proposition reveals.

Proposition 3.2. *Assume selective influence. The independent parallel exhaustive processing predicts underadditivity in the survivor function when an even number of channels are processed, and predicts overadditivity when an odd number of channels are processed.*

Proof. Without the consideration of base time, the total reaction time in parallel exhaustive process models should be the maximum of the reaction time on each process, i.e., $T = \max(T_{X1}, \ldots, T_{Xn})$. Thus, we have

$$
\begin{aligned}
\mathrm{SIC}^n_{\mathrm{par.AND}}(t) &= \Delta^n_{X_1,\ldots,X_n} P(\max(T_{X1}, \ldots, T_{Xn}) > t) \\
&= \Delta^n_{X_1,\ldots,X_n}[1 - P(\max(T_{X1}, \ldots, T_{Xn}) \leqslant t)] \\
&= -\Delta^n_{X_1,\ldots,X_n} P(\max(T_{X1}, \ldots, T_{Xn}) \leqslant t)
\end{aligned}
$$

$$
\begin{aligned}
&= -\Delta^n_{X_1,\ldots,X_n}[P(T_{X1} \leqslant t) \times \cdots \times P(T_{Xn} \leqslant t)] \\
&\quad \text{(factoring)} \\
&= -[P(T_{XnL} \leqslant t) - P(T_{XnH} \leqslant t)] \\
&\quad \times \Delta^{n-1}_{X_1,\ldots,X_{n-1}}[P(T_{X1} \leqslant t) \times \cdots \times P(T_{X(n-1)} \leqslant t)] \\
&= [F_{nL}(t) - F_{nH}(t)] \times \mathrm{SIC}^{n-1}_{\mathrm{par.AND}}(t). \quad (13)
\end{aligned}
$$

Because of the action of selective influence on the CDF, $F_{nL}(t) < F_{nH}(t)$ for all t. As we know that $\mathrm{SIC}^2_{\mathrm{par.AND}}(t)$ is negative, we can infer that $\mathrm{SIC}^n_{\mathrm{par.AND}}(t)$ is negative when n is even and is positive when n is odd. □

This flip-flopping of the SIC function according to whether n is odd or even, and only for the exhaustive stopping rule, is quite surprising and should be highly useful in identifying mental architectures.

We next resume our investigation of serial systems with the same two stopping rules; minimum time first.

3.3. Serial minimum time processing

Recall that in the case of serial processing, a minimum time stopping rule is simplicity itself − the very first completion brings processing to a halt. This simplicity, reinforced with selective influence, extracts a strong result.

Proposition 3.3. *Assume selective influence. The independent serial minimum time processing predicts that the SIC will always be zero as a function of time t, for every n.*

Proof. The total reaction time in serial minimum time process models should be the reaction time of the process chosen to be the first one processed. Let p_i denote the probability of processing the ith item first. Then the survivor function on the total reaction time can be expressed as

$$
\begin{aligned}
S(t) &= P(T > t) \\
&= \sum_{i=1}^n P(T > t | T = T_{Xi}) \times P(T = T_{Xi}) \\
&= \sum_{i=1}^n p_i \times P(T_{Xi} > t). \quad (14)
\end{aligned}
$$

Due to the complete factorial design, the n-process SIC function is composed of 2^n survivor functions, and all items will cancel each other out in the ensuing contrast function. Thus, the minimum time stopping rule leads immediately to a decisive, and invariant prediction. □

3.4. Serial maximum time processing

One of the most intriguing architectural signatures of the SIC curve when $n = 2$ is that of serial exhaustive processing, exhibiting as it does the presence of wiggles above and below zero. As we would expect, for arbitrary n, the mean interaction contrast for a serial process with an exhaustive stopping rule must be 0 due to the additivity of the processing times along with selective influence. The following proposition confirms this result using survivor functions but it also follows almost immediately from the foregoing statement.

Proposition 3.4. *Assume selective influence. The independent serial exhaustive processing predicts that the integral of the SIC function is always equal to zero, for every n.*

Proof. Without the consideration of base time, the total reaction time in serial exhaustive process models should be the sum of the reaction time on each process, i.e., $T = T_{X1} + \cdots + T_{Xn}$. We reveal

the relationship between the n-stage SIC function and the $(n-1)$-stage SIC function as below

$$
\begin{aligned}
\mathrm{SIC}^n_{\mathrm{ser.AND}}(t) &= \Delta^n_{X_1,\dots,X_n} P(T_{X1} + \cdots + T_{Xn} > t) \\
&= -\Delta^n_{X_1,\dots,X_n} P(T_{X1} + \cdots + T_{Xn} \leqslant t) \\
&= -\Delta^n_{X_1,\dots,X_n}\left[\int_0^t [f_n(t_n) \right.\\
&\qquad \left. \times P(T_{X1} + \cdots + T_{X(n-1)} \leqslant t - t_n)]dt_n\right] \\
&= -\int_0^t \Delta^n_{X_1,\dots,X_n}\Big[[f_n(t_n) \\
&\qquad \times P(T_{X1} + \cdots + T_{X(n-1)} \leqslant t - t_n)]\Big]dt_n \\
&\qquad \text{(factoring)} \\
&= -\int_0^t \Big[[f_{nL}(t_n) - f_{nH}(t_n)] \\
&\qquad \times [\Delta^{n-1}_{X_1,\dots,X_{n-1}} P(T_{X1} + \cdots + T_{X(n-1)} \leqslant t - t_n)]\Big]dt_n \\
&= -\int_0^t \Big[[f_{nH}(t_n) - f_{nL}(t_n)] \\
&\qquad \times \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t - t_n)\Big]dt_n.
\end{aligned}
\tag{15}
$$

Notice that $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ is a convolution. By Fubini's theorem, the integral of the convolution of two functions on the whole space is simply obtained as the product of the integrals of each function, thus the integral of the SIC function on the whole space could be rewritten as

$$
\begin{aligned}
\int_0^\infty \mathrm{SIC}^n_{\mathrm{ser.AND}}(t)dt &= -\int_0^\infty \left[\int_0^t [f_{nH}(t_n) - f_{nL}(t_n)] \right.\\
&\qquad \left. \times \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t - t_n)dt_n\right]dt \\
&= -\int_0^\infty [f_{nH}(t) - f_{nL}(t)]dt \\
&\qquad \times \int_0^\infty \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)dt.
\end{aligned}
\tag{16}
$$

Since both $f_{nH}(t)$ and $f_{nL}(t)$ are density functions, the first integral of the above equation is always zero. Thus for every n, the integral of the SIC function of serial exhaustive model is predicted to be zero. □

The next proposition indicates that the SIC function will flip over when the number of processes changes from even to odd and vice versa.

Proposition 3.5. *Assume selective influence. The independent serial exhaustive processing predicts that the SIC function must be negative for small times if n is even, and it must be positive for small times if n is odd.*

Proof. It is easy to confirm this statement when we examine the relationship between $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ and $\mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)$ shown in Eq. (15). Notice that the difference between the two density functions should be positive for small t because of selective influence at the density crossing level. Thus, when t is small, $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ and $\mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)$ have different signs, because of the negative sign before the integral in the above equation. □

Propositions 3.4 and 3.5 provide us some key properties of the SIC functions for serial exhaustive models, from which we garner some idea about how the SIC signatures behave, and how they assist in identification of the architectures and stopping rules. With regard to the further analysis of SIC curves, recall that in the previous section, we discovered that in the $n = 2$ case, a mild assumption of log-concavity guarantees that there is only a single wiggle through zero. Thus an intriguing question asks how many zeros there will be in $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$? Is there any trend we can discover as n increases? The following discussion shows that the answer to all these questions depends on the analytic character of the function. For functions of certain form, the zeros of $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ will increase by 1 as n increases by 1.

Next, it may be remembered that in the previous section we determined the zeros of $\mathrm{SIC}^2_{\mathrm{ser.AND}}(t)$ via the shape of its integral. Here, we again explore the integral of $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ for the same purpose. Let $I^n(s)$ denote the integral of $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ with variable upper limit. Eq. (17) reveals the relationship between $I^n(s)$ and $I^{n-1}(s)$, which eventually has the same convolution form as the relationship between $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ and $\mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)$.

$$
\begin{aligned}
I^n(s) &= -\int_0^s \left[\int_0^t [f_{nH}(t_n) - f_{nL}(t_n)] \times \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t - t_n)dt_n\right]dt \\
&\qquad \text{(Change the order of integration)} \\
&= -\int_0^s \left[\int_{t_n}^s [f_{nH}(t_n) - f_{nL}(t_n)] \times \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t - t_n)dt\right]dt_n \\
&= -\int_0^s \left[[f_{nH}(t_n) - f_{nL}(t_n)] \times \int_{t_n}^s \mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t - t_n)dt\right]dt_n \\
&= -\int_0^s \left[[f_{nH}(t_n) - f_{nL}(t_n)] \times I^{n-1}(s - t_n)\right]dt_n.
\end{aligned}
\tag{17}
$$

Recall that with the assumption of log-concavity, $I^2(s)$ is a unimodal function. So following the proof in Proposition 2.2, we assert that $I^3(s)$ is a one-wiggle S-shape function, just as $\mathrm{SIC}^2_{\mathrm{ser.AND}}(t)$, if $F_{X3H}(t) - F_{X3L}(t)$ is strictly log-concave. More generally, with the assumption of strictly log-concavity of DF_{Xi}, we can infer that $I^n(s)$ should have the same key signature as $\mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)$ by simple math induction. Now the problem of convolution converts to the operation of taking the derivative, that is, how the number of 0-crossings change from $\mathrm{SIC}^{n-1}_{\mathrm{ser.AND}}(t)$ to $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ converts to the question of how that number changes when we take the derivative of $I^n(s)$. Unfortunately, there appears to be no uniform answer for that question. In fact, for some particular functions such as the Gaussian density function the outcome is pleasant since the nth derivative would have n zeros, which leads to the result that zeros of $\mathrm{SIC}^n_{\mathrm{ser.AND}}(t)$ will increase by 1 as n increases by 1.

However, in the general case, the question of the number of zeros of the derivatives could be a quite complicated issue due to the flexible forms of the distributions. As a result, although our hypothesis that zeros would increase by 1 as n increases has been observed in simulations, theoretical analysis suggests that this hypothesis could be violated, especially when n is a large number.

Fig. 4 exhibits simulation results which document the key signatures of the varying SIC functions. Briefly, Fig. 4 first depicted the fact that parallel minimum time processing models always predict an entirely positive SIC function (first column). Secondly, the curious flip flopping of the parallel exhaustive systems from positive to negative as n goes from odd to even is shown in the second column. Thirdly, the absolute invariance of the minimum time serial processing models' prediction that the entire curve is identical to 0 appears in the third column of panels. Finally, the fascinating multiplying of the wiggles in the serial exhaustive systems as n grows is illustrated in the fourth column of Fig. 4.
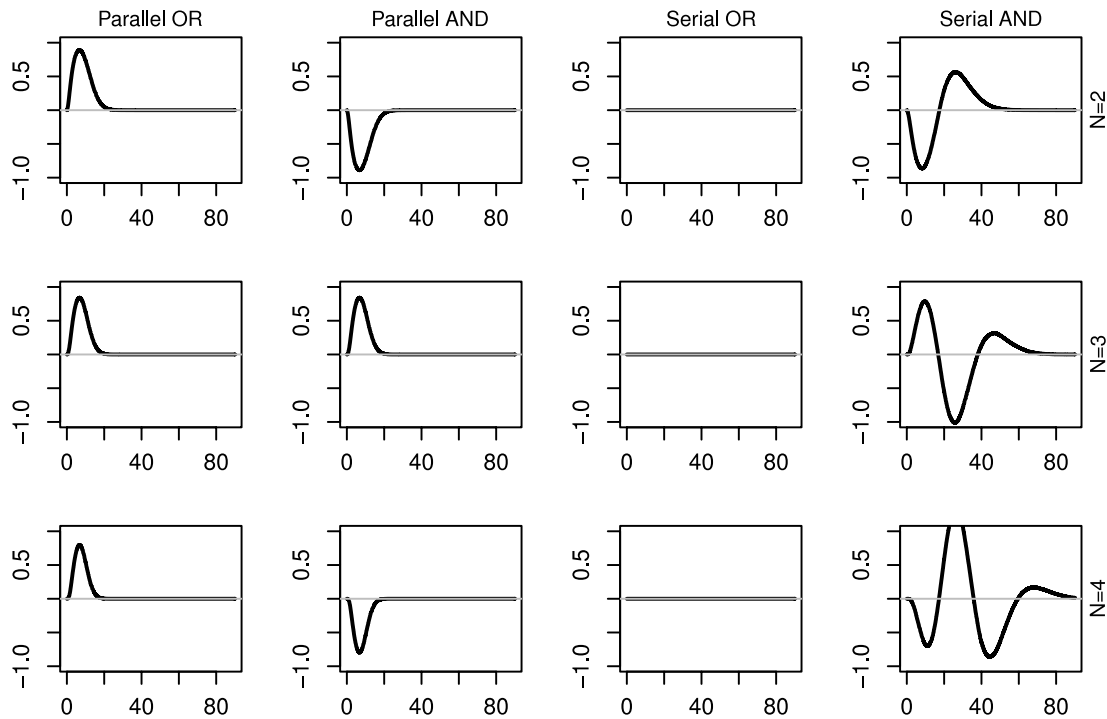
Fig. 4. Predicted survivor interaction contrast forms for Parallel-OR (first column), Parallel-AND (second column), Serial-OR (third column) and Serial-AND (fourth column) processing models, for varying n from 2 to 4 (top to bottom). Simulations are based on Gamma distribution.

4. Exemplary experiment

In a typical short-term memory search task, a small set of items is quickly memorized, and then it is followed by the presentation of a target. Participants are required to indicate, by pressing either the 'yes' or 'no' button, whether the target was in the memory set. The major interest of the task is to see whether the comparisons of target and memory set items were carried out in serial or in parallel. The data we present below come from an earlier short-term memory search study in which the size of the memory set was manipulated ($n = 2$ or 4). The details of the experiment design and the results for $n = 2$ condition have been reported in Townsend and Fifić (2004), in which they revealed that different individuals may evince either parallel or serial processing. In this paper we take one participant whose data in the $n = 2$ condition strongly support serial processing for example, and we focus on his data in the $n = 4$ condition. With a larger memory set of size 4, we might expect to see more and consistent evidence to support serial processing. The $n = 4$ design also allows for examination of wiggle behavior of the observed SIC functions.

4.1. Method

The task was carried out in Belgrade, using stimuli crafted from Serbian linguistic features. Five participants were recruited for the experiment. Each participant participated in both the $n = 2$ and $n = 4$ conditions. A total of 12 stimuli were constructed and they were all pseudo Serbian words in consonant–vowel–consonant form. In the experiment, each trial consisted of a fixation point and warning low-pitch tone for 1 s, successive presentation of two or four items in the memory set for 1200 ms, an interstimulus interval (ISI), and a target. Participants were required to indicate, by pressing either the 'yes' or 'no' button, whether the target was in the memory set or not. The ISI was also manipulated, ISI = 700 or 2,000 ms. In the $n = 4$ task, participants were run for 44 blocks of 128 trials for each ISI condition, and they were required to respond as fast as they could without compromising accuracy. Target items

and memory-set items were manipulated across trials to produce low and high levels of item-target dissimilarity. For examples of stimuli and more details on the experimental setup, see Townsend and Fifić (2004).

4.2. Data analysis

Our analyses focused only on target-absent trials. As in Townsend and Fifić (2004), the experimental factor for each distractor item (which coincided with a stage in serial processing or a channel in parallel processing) was its similarity to the target. This type of manipulation should permit selective influence on exhaustive ('no') trials but could introduce confounds on single-target, self-terminating ('yes') trials. Naturally, on target-absent trials, participants must exhaustively process all four comparisons of target and memory set items to guarantee a correct response.

To implement the double factorial paradigm, we conducted two levels of item-target dissimilarity. High dissimilarity was designated as the high salience condition and low dissimilarity was designated as the low salience condition. Thus the factors of interests were positions of item in the set (1, 2, 3, 4) × phonemic similarity (low, high). The factorial combination of four factors with their two salience levels leads to sixteen experimental conditions ($2 \times 2 \times 2 \times 2$). For instance, LHHH indicates a condition in which the first factor is of low salience and all other factors are of high salience. Simpler embedded factorial designs can be derived from the full design by reducing the number of factors. For instance, from the four-target design it is possible to obtain 4 three-target designs, by dropping out any of the positional factors. Thus, the following will be obtained: x234, 1x34, 12x4 and 123x. The 'x' denotes position that was dropped out from the analysis. Similarly, excluding yet another factor would derive 6 two-target designs: xx34, x2x4, x23x, 1xx4, 1x3x, 12xx. In other words, merging selected factors allow us to analyze the underlying architecture of any two, three or four mental processes within one experiment.

Townsend and Fifić (2004) working with $n = 2$, discovered individual differences and, for some participants, the influence of

Table 1
Results of the Kolmogorov–Smirnov test applied to the full task.

	hhhh	lhhh	hlhh	hhlh	hhhl	llhh	lhlh	lhhl	hllh	hlhl	hhll	lllh	llhl	lhll	hlll	llll
hhhh		>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***
lhhh	−					>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***
hlhh	−					>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***
hhlh	−					>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***
hhhl	−					>***	>***	>***	>***	>***	>***	>***	>***	>***	>***	>***
llhh	−	−	−	−	−							>***	>***	>***	>***	>***
lhlh	−	−	−	−	−							>***	>***	>**	>***	>***
lhhl	−	−	−	−	−							>***	>***	>***	>***	>***
hllh	−	−	−	−	−							>***	>***	>**	>***	>***
hlhl	−	−	−	−	−							>***	>***	>***	>***	>***
hhll	−	−	−	−	−							>***	>***	>***	>***	>***
lllh	−	−	−	−	−	−	−	−	−	−	−					>***
llhl	−	−	−	−	−	−	−	−	−	−	−					>***
lhll	−	−	−	−	−	−	−	−	−	−	−					>***
hlll	−	−	−	−	−	−	−	−	−	−	−					>***
llll	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$
$>^{***}$: $\max(F_{HHHH}(t) - F_{LHHH}(t))$ is significant, $p < 0.01$, etc.
$-$: $\max(F_{LHHH}(t) - F_{HHHH}(t))$ is not significant, etc.

the ISI between memory-set presentation and the probe. However, all the SIC functions decisively adhered to either a serial or parallel form. To demonstrate the use of the new theories for revealing architecture for multiple processes, we take one participant (Participant 1 in Townsend and Fifić (2004)) whose data consistently support serial processing as an example of our new theoretical results. We were therefore curious to examine behavior for larger n, particularly since most memory search experiments are based on $n > 2$. We were also keen to test the within-architecture predictions concerning behavior as the number of processes altered across $n = 2, 3, 4$, utilizing the above 'factor reduction' strategy.

Evidence of selective influence was first assessed by a two-sample Kolmogorov–Smirnov test. As noted earlier, selective influence implies that the reaction time when the target is low salience is stochastically larger than the reaction time when the target is high salience. In other words, it implies that the distribution functions of RT should exhibit a stochastic ordering across conditions. In the two-sample Kolmogorov–Smirnov test, the null hypothesis is that the two empirical distributions are equal, and the alternative hypothesis is that the two distributions are different, in the sense that the maximum of the difference between two distributions is significantly large. To test for illegitimate cumulative distribution function crossings (logically equivalent to crossings of the survivor functions), we applied the one-sided Kolmogorov–Smirnov test twice on each pair of conditions to assess selective influence as is common practice (e.g., Townsend & Nozawa, 1995). For instance, when assessing the selective influence between the conditions HHHH and HHHL, we require that $\max(F_{HHHH} - F_{HHHL})$ is significant and $\max(F_{HHHL} - F_{HHHH})$ is not significant.

Next, the architecture and stopping rule first assessed by the adjusted rank transform (ART; Sawilowsky (1990)) test. The ART test is a nonparametric test for an interaction between two discrete variables. We apply it to distinguish serial versus parallel models. The null hypothesis is that there is no interaction between the salience manipulations on each process, so rejecting the null hypothesis falsifies serial models (MIC = 0). The architecture and stopping rule were additionally examined by a nonparametric statistical test of SIC, which was recently developed by Houpt and Townsend (2010). This test performs two separate null hypothesis tests with a D-statistic: One is for testing whether the empirical

SIC function (ESIC) has a significant positive deviation from zero, and the other is pertinent to testing whether the ESIC has a significant negative deviation from zero. Referring to Fig. 1, rejecting the null hypothesis of the first test falsifies serial-OR and parallel-AND models, whereas rejecting the null hypothesis of the second test falsifies serial-OR and parallel-OR models. It may be observed that the significance test of SIC function cannot presently distinguish a serial-AND model from a coactive model, since both models are expected to have a significant positive deviation and a significant negative deviation. Thus combining the statistical test of SIC with the ART test is necessary. For more details of the statistical test of SIC, see Houpt and Townsend (2010).

4.3. Result

Selective influence. The Kolmogorov–Smirnov tests were applied to each pair of high versus low salience conditions in the full task. Table 1 shows the K–S results for the participant: As shown in Table 1, in every case where an order was predicted, the appropriate ordering was supported, either at the $p = 0.01$ or $p = 0.05$ level. Additionally, non-significant results were found for each pair of conditions with the application of the inverse direction.

Architecture and stopping rule. The empirical SIC (ESIC) functions for the same participant are plotted in Fig. 5. The top figure displays the four-target ESIC function ($n = 4$), the middle row shows four reduced three-target ESICs ($n = 3$), and the bottom row of the Fig. 5 shows a total of 6 two-target ESICs ($n = 2$). The statistical tests were applied to analyze the $n = 2$ ESIC functions, with results summarized in Table 2. The first column in Table 2 shows the statistical results of the ART test, and the rest of the table exhibits the maximum deviations from the statistical test of ESIC.

As can be seen in the last column of Table 2, all of the six models revealed a significant deviation of the ESIC ($p < 0.01$). Hence, we can reject the null hypothesis that the true SIC is flat. Furthermore, five of the six models exhibited a non-significant result in their ART test. This result provides evidence supporting serial models. This observation, together with the significant result of ESIC, provides strong evidence in favor of serial-AND processing that is consistent with our hypothesis. There is only one exception, the model x2x4,
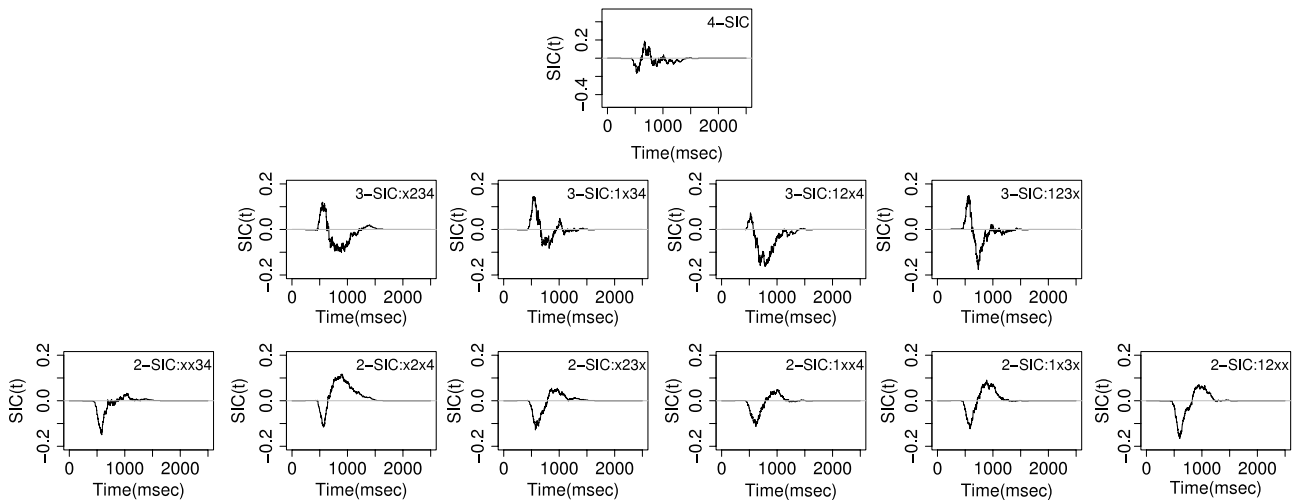
Fig. 5. ESIC(t) for one participant. The top row is the ESIC for the original 4-target task. The middle and bottom rows are the ESICs estimated from the reduced 3-target and 2-target tasks, respectively.

Table 2
Results of the SIC and the ART test statistics applied to the data from the reduced 2-target tasks.

Model	ART	D^+	D^-	$D = \max(D^+, D^-)$
xx34	0.013	0.033	0.150***	0.150***
x2x4	13.001***	0.117***	0.116***	0.117***
1xx4	0.053	0.048	0.113***	0.113***
x23x	0.894	0.056	0.128***	0.128***
1x3x	1.801	0.091	0.123***	0.123***
12xx	0.046	0.071*	0.166***	0.166***

* $p < 0.1$ ** $p < 0.05$.
*** $p < 0.01$.

exhibiting a significant result of the ART test, which implies a possible interaction between processes 2 and 4. Since both D^+ and D^- are significant from zero ($D^+ = 0.117$, $D^- = 0.116$, $p < 0.01$), unfortunately we cannot rule out the coactive model as a model of processing in this case (see Fig. 1, right panel).

The current SIC statistical tests for significance are limited by their confinement to $n = 2$. For $n > 2$, however, even without statistical tools, we can nevertheless assay the qualitative form of the ESIC signatures. In fact, as we can see, the signatures of ESIC when $n = 3$ exhibits clear M-shaped patterns. Each ESIC was positive for early processing times and negative in the middle, and go up again for the later processing times. Those M-shaped patterns provide preliminary evidence for a three-stage serial AND processing. The top 4-target ESIC function appears to cross the x-axis at least twice although crossings at the tail are rather minuscule. The 4-target ESIC function does start negative, and then becomes positive, and then appears to wiggle.

Moreover, for each of the ESIC functions we observed equal positive and negative areas, confirming the prediction of serial AND model. Intriguingly, the ESIC functions further appear to obey the flip-flopping property when the number of processes changes from even to odd. Thus we tentatively infer that this participant searches her/his short-term memory serially for up to 4 memory items.

5. Summary and discussion

This paper expands the knowledge base concerned with identification of the simple but crucial architectures referred to as serial and parallel processing along with two major decisional stopping rules. As in almost all of our developments on these topics, the predictions have been entirely distribution free. This attribute renders the related diagnostic tests more robust and permits the testing

(e.g., falsification) of entire classes of models. The present developments continue that tradition.

The SIC behavior of parallel processing in league with the major stopping rules and for $n = 2$, has been definitively established since 1995 (Townsend & Nozawa, 1995). Likewise, minimum processing time serial signatures were already known for $n = 2$. However, in the important case of serial exhaustive processing, the exact form of the SIC signature has been unclear, even for $n = 2$. The early results showed that it must start negative (i.e., for small time values, t), express positive areas as well and that the positive plus negative areas must equal zero. Many simulations and some experimental data (Fific, Little, & Nosofsky, 2010; Fific, Nosofsky, & Townsend, 2008; Little, Nosofsky, & Denton, 2011; Townsend & Fific, 2004) further have suggested a single 0-crossing but this property had not been proven.

The present investigation thus began by exploring the precise behavior of the serial exhaustive SIC function for $n = 2$. We found that a single wiggle (i.e., one 0-crossing) cannot be absolutely guaranteed since there exist probability distributions that do not elicit that property. However, it was proven that: A. There must be an odd number of crossings for any distributions. B. A rather mild condition known as log-concavity is sufficient as a guarantor of a single 0-crossing.

Up until now, there has further been a lack of knowledge concerning how the architectural signatures act when n is varied. The second major part of the investigation pursued this issue. The present theoretical analysis provides this information for parallel and serial models in conjunction with both the minimum time and maximum time stopping rules. Although in some cases the exact shape of the SIC curve is still unknown (e.g., the SIC for the serial exhaustive model for $n > 2$) due to the flexibility of distribution forms of the processing time, we have provided thorough analysis specifying critical and differing traits of SIC functions under different models that should prove vital in identifying fascinating these elementary, but fundamental, architectures along with their stopping rules.

The present developments significantly expand the arena within which SFT and RT data can assay elementary architectures and stopping rules and we anticipate their application in basic and applied research.

References

Atkinson, R., Holmgren, J., & Juola, J. (1969). Processing time as influenced by the number of elements in a visual display. *Attention, Perception, & Psychophysics*, *6*, 321–326.

Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology (vol. 2)*. New York: Wiley.

Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979.

Chung, K. L. (1953). Sur les lois de probabilites unimodales. *Comptes Rendus de l'Academie des Sciences*, *236*, 583–584.

Dharmadhikari, S. W., & Joag-Dev, K. (1988). *Unimodality, convexity, and applications*. Academic Press.

Donders, F. C. (1868). Over de snelheid van psychische processen. Onderzoekingen gedaan in het physiologisch laboratorium der Utrechtsche hoogeschool, 1868–1869. *Tweeds Reeks*, *II*, 92–120.

Dzhafarov, E. N. (1997). Process representations and decompositions of response times. In A. A. J. Marley (Ed.), *Choice, decision and measurement: essays in honor of R. Duncan Luce*. Mahwah, NJ: Erlbaum.

Dzhafarov, E. N. (1999). Conditionally selective dependence of random variables on external factors. *Journal of Mathematical Psychology*, *43*, 123–157.

Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrica*, *68*, 7–25.

Dzhafarov, E. N., Schweickert, R., & Sung, K. (2004). Mental architectures with selectively influenced but stochastically interdependent components. *Journal of Mathematical Psychology*, *48*, 51–64.

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Attention, Perception, & Psychophysics*, *1*, 245–252.

Eriksen, C. W., & Spencer, T. R. (1969). Rate of information processing in visual perception: some results and methodological considerations. *Journal of Experimental Psychology Monographs*, *79*, 1–16.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*(3), 145–154.

Estes, W. K. (1959). The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: a study of a science (vol. 2)*. New York: McGraw-Hill.

Estes, W. K., & Taylor, H. A. (1964). A detection method and probabilistic models for assessing information processing from brief visual displays. *Proceedings of the National Academy of Science*, *52*, 446–454.

Estes, W. K., & Taylor, H. A. (1966). Visual detection in relation to display size and redundancy of critical elements. *Perception and Psychophysics*, *1*, 9–16.

Estes, W. K., & Wessel, D. L. (1966). Reaction time in relation to display size and correctness of response in forced-choice visual signal detection. *Perception and Psychophysics*, *1*, 369–373.

Fific, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: a synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review*, *117*, 309–348.

Fific, M., Nosofsky, R. M., & Townsend, J. T. (2008). Information-processing architectures in multidimensional classification: a validation test of the systems factorial technology. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 356–375.

Friedman, M. P., Burke, C. J., Cole, M., Keller, L., Millward, R. B., & Estes, W. K. (1964). Two-choice behavior under extended training with shifting probabilities of reinforcement. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Stanford, CA: Stanford University Press.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Huntington, N.Y.: Krieger.

Haber, R. N., & Hershenson, M. (1973). *The psychology of visual perception*. Oxford, England: Holt, Rinehart & Winston.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11–36.

Houpt, J. W., & Townsend, J. T. (2010). The statistical properties of the survivor interaction contrast. *Journal of Mathematical Psychology*, *54*, 446–453.

Hull, C. L. (1952). *A behavior system: an introduction to behavior theory concerning the individual organism*. New Haven, CT, US: Yale University Press.

Ibragimov, I. A. (1956). On the composition of unimodal distributions. *Theory of Probability and its Applications*, *1*, 255–360.

Koch, S. (1954). Clark l. Hull. In W. K. Estes, S. Koch, K. MacCorquodale, P. E. Meehl, C. G. Mueller, W. N. Schoenfeld, & W. S. Verplanck (Eds.), *Modern learning theory*. East Norwalk, CT, US: Appleton-Century-Crofts.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement (vol. 1.)*. New York: Academic Press.

Kujala, J. V., & Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, *52*, 128–144.

Little, D. R., Nosofsky, R. M., & Denton, S. E. (2011). Response-time tests of logical-rule models of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1–27.

Murdock, B. B. (1971). A parallel-processing model for scanning. *Attention, Perception, & Psychophysics*, *10*, 289–291.

Nickerson, R. S. (1972). Binary-classification reaction time: a review of some studies of human information-processing capabilities. *Psychonomic Monograph Supplements*, *4*, 275–318.

Posner, M. I. (1978). *Chronometric explorations of mind*. NJ: Erlbaum: Englewood Heights.

Roberts, P., & Zinnes, J. L. (1963). In R. R. B. R. D. Luce, & E. Galanter (Eds.), *Handbook of mathematical psychology (vol. 1)*. New York: Wiley.

Sawilowsky, S. S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research*, *60*, 91–126.

Scharff, A., Palmer, J., & Moore, C. M. (2011). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 813–833.

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

Schweickert, R. (1982). The bias of an estimate of coupled slack in stochastic pert networks. *Journal of Mathematical Psychology*, *26*, 1–12.

Schweickert, R., & Giorgini, M. (1999). Response time distributions: some simple effects of factors selectively influencing mental processes. *Psychonomic Bulletin & Review*, *6*, 269–288.

Schweickert, R., Giorgini, M., & Dzhafarov, E. N. (2000). Selective influence and response time cumulative distribution functions in serial-parallel task networks. *Journal of Mathematical Psychology*, *44*, 504–535.

Schweickert, R., & Townsend, J. T. (1989). A trichotomy interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (pert) networks. *Journal of Mathematical Psychology*, *33*, 328–347.

Schweickert, R., & Wang, Z. (1993). Effects on response time of factors selectively influencing processing in acyclic task networks with or gates. *British Journal of Mathematical and Statistical Psychology*, *46*, 1–30.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs General and Applied*, *74*, 1–29.

Sperling, G. (1967). Successive approximations to a model for short-term memory. In A. Sanders (Ed.), *Attention and performance*. Amsterdam: North- Holland.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652.

Sternberg, S. (1969). The discovery of processing stages: extensions of Donders' method. *Acta Psychologica*, *30*, 276–315.

Sternberg, S. (1975). Memory scanning: new findings and current controversies. *Quarterly Journal of Experimental Psychology*, *27*, 1–32.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*, 401–409.

Townsend, J.T. (1969). Stochastic representations of serial and parallel processes. In *Paper presented at Midwest Mathematical Psychology Meetings, Indiana University*.

Townsend, J. T. (1971). A note on the identifiability of parallel and serial processes. *Attention, Perception, & Psychophysics*, *10*, 161–163.

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

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. Kantowitz (Ed.), *Human information processing*. Hillsdale, NJ: Erlbaum Press.

Townsend, J. T. (1976). Serial and within-stage independent parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology*, *14*, 219–238.

Townsend, J. T. (1981). Some characteristics of visual whole report behavior. *Acta Psychologica*, *47*, 149–173.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, *28*, 363–400.

Townsend, J. T. (1990a). Serial vs. parallel processing: sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46–54.

Townsend, J. T. (1990b). Truth and consequences of ordinal differences in statistical distributions: toward a theory of hierarchical inference. *Psychological Bulletin*, *108*, 551–567.

Townsend, J. T. (1992). Don't be fazed by phaser: beginning exploration of a cyclical motivational system. *Behavior Research Methods, Instruments, & Computers*, *24*, 219–227.

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

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

Townsend, J. T., Fific, M., & Neufeld, R. W. J. (2007). Assessment of mental architecture in clinical/cognitive research. In T. A. Treat, R. R. Bootzin, & T. B. Baker (Eds.), *Psychological clinical science: papers in honor of Richard M. Mcfall*. Mahwah, NJ: Lawrence Erlbaum Associates.

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*, 321–359.

Townsend, J. T., & Thomas, R. (1994). Stochastic dependencies in parallel and serial models: effects on systems factorial interactions. *Journal of Mathematical Psychology*, *38*, 1–34.

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

Townsend, J. T., Yang, H., & Burns, D. M. (2011). Experimental discrimination of the world's simplest and most antipodal models: the parallel-serial issue. In H. Colonius, & E. N. Dzhafarov (Eds.), *Descriptive and normative approaches to human behavior in the advanced series on mathematical psychology*. Singapore: World Scientific.

Vorberg, D. (1977). On the equivalence of parallel and serial models of information processing. In *Paper presented at the tenth annual mathematical psychological meeting*.