

10-1-2010

The Statistical Properties of the Survivor Interaction Contrast

Joseph W. Houpt

Wright State University - Main Campus, joseph.houpt@wright.edu

James T. Townsend

Follow this and additional works at: <http://corescholar.libraries.wright.edu/psychology>

 Part of the [Cognition and Perception Commons](#), [Cognitive Psychology Commons](#), [Quantitative Psychology Commons](#), and the [Statistical Models Commons](#)

Repository Citation

Houpt, J. W., & Townsend, J. T. (2010). The Statistical Properties of the Survivor Interaction Contrast. *Journal of Mathematical Psychology*, 446-453.

<http://corescholar.libraries.wright.edu/psychology/12>

This Article is brought to you for free and open access by the Psychology at CORE Scholar. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact jane.wildermuth@wright.edu, riley.stoermer@wright.edu.

The Statistical Properties of the Survivor Interaction Contrast

Joseph W. Houpt^{a,*}, James T. Townsend^a

^aDepartment of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA

Abstract

The Survivor Interaction Contrast (SIC) is a powerful tool for assessing the architecture and stopping rule of a model of mental processes (Townsend and Nozawa, 1995). Despite its demonstrated utility, the methodology has lacked a method for statistical testing until now. In this paper we briefly describe the SIC then develop some basic statistical properties of the measure. These developments lead to a statistical test for rejecting certain classes of models based on the SIC. We verify these tests using simulated data, then demonstrate their use on data from a simple cognitive task.

Key words: Mental Architecture, Human Information Processing, Survivor Interaction Contrast, Nonparametric

Much of scientific psychology and cognitive science can be viewed as seeking to understand the mechanisms and dynamics of perception, thought, and action. Two processing attributes that have been of particular interest to psychologists are the architecture, or temporal relationships between sub-processes of the system, and stopping rule, which dictates how many of the sub-processes must be completed for the system to finish. The Survivor Interaction Contrast (SIC) is a powerful tool for assessing the architecture and stopping rule of a model of a mental process (Townsend and Nozawa, 1995). Despite its demonstrated utility, the methodology has lacked a method for statistical testing. In this paper we briefly describe the SIC then explore some basic statistical properties of the measure. These developments lead to a statistical test for rejecting certain classes of models based on the SIC.

Determining the architecture of a mental process is vital to understanding how that process works. In particular, any relatively detailed model of a mental process must make an assumption about the temporal relationships between the components of that process. For example, if there is a set of items in memory to compare against a target, each item may be compared to the target sequentially, that is serially, such as depicted at the top of Figure 1. Or it may be that all items are compared against the stimuli at the same time, in parallel, as depicted in the center of Figure 1. The question of architecture arises even in simple cognitive tasks. Consider the recent experiment by Eidels et al. (2010) in which a dot was presented to a participant's right or left visual field, both, or neither and the participant's task was to decide whether or not a dot was shown. The left and right visual fields are likely processed in parallel, although it is important to rule out

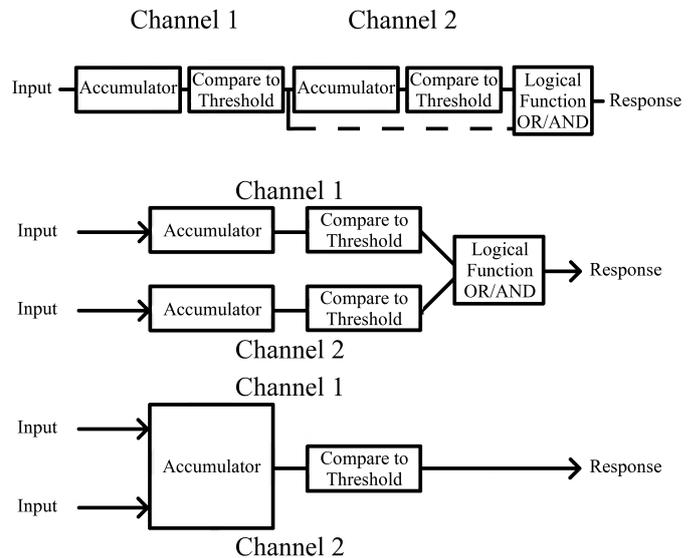


Figure 1: Diagrams of serial, parallel and coactive information processing architectures.

a serial architecture. Another alternative is that, as the information from the left and right visual fields is accumulated, the information is pooled, depicted at the bottom of Figure 1. Processes in which the information from each parallel sub-process is pooled toward a single decision are *coactive*.

These three general classes of architecture, serial, parallel, and coactive, do not span the entire range of possible configurations of mental sub-processes. Some blend of serial and parallel could also be possible, where a subset of the sub-processes occurs in parallel and another subset occurs in serial. Although we limit our scope in this paper to models of these most frequently studied cases, there has

*Corresponding author

Email addresses: jhoupt@indiana.edu (Joseph W. Houpt), jtownsen@indiana.edu (James T. Townsend)

been significant theoretical work on more general architectures, such as the PERT networks (Schweickert, 1978; Schweickert et al., 1989).

Despite the clear structural differences between serial, parallel and coactive architectures, distinguishing between them based on observable data can be challenging. One reason for the difficulty is that, in general, a serial architecture can mimic the response time patterns of a parallel architecture even at the level of the distribution. In many cases a parallel system can mimic a serial process as well (Townsend, 1972, 1974; Townsend and Ashby, 1983, Chapter 14).

Consider the Sternberg paradigm as an example (Sternberg, 1966). In this paradigm a participant must search through memory for a target among a varying number of distractors. The claim has often been made, including by Sternberg (1966), that if the mean response time increases linearly with the number of distractors, then each item must be processed in serial. Suppose all items must be examined (the so-called exhaustive stopping rule see below for more on this topic). Then ordinary serial processing does indeed predict a straight increasing RT line. In contrast, most parallel models do not make this prediction, as Sternberg (1966) indicated. However, consider a parallel model with a fixed amount of resources divided among the items to be processed. Suppose also that as each item is completed, the attentional resources devoted to the finished item is reallocated to the remaining, unfinished items. This model has been called the fixed capacity, reallocation parallel model (Townsend and Ashby, 1983, Chapter 4). The more items there are, the fewer resources each individual item has, thereby reducing the rate of each sub-process, but the overall production rate remains constant within a trial, just as serial processing predicts Townsend (1990). Of course, these verbal comments can be made rigorous. For instance, the above mimicking results can be reproduced within distributions based on exponential intercompletion times (the duration between successive item completions; e.g., Townsend, 1972). Many other parallel and serial models can mimic each other in the strong sense of distributional equivalence for arbitrary distributions (see Townsend, 1976; Townsend and Ashby, 1983, Chapter 14). The statistics developed in this study are associated with theory-driven experimental designs which, under the specified assumptions, do not permit model mimicking.

Another fundamental property of mental processes is the stopping rule for the system. The stopping rule determines the number of completed sub-process necessary for the system to finish. For example, a system with two sub-processes could finish when either one of the sub-processes finish or may need both sub-process to complete. If a system is finished as soon as the first sub-process completes, it has a first-terminating (OR) stopping rule. A system that requires all of the sub-processes to complete before terminating has an exhaustive (AND) stopping rule. The stopping rule of a process can often be dictated by the task;

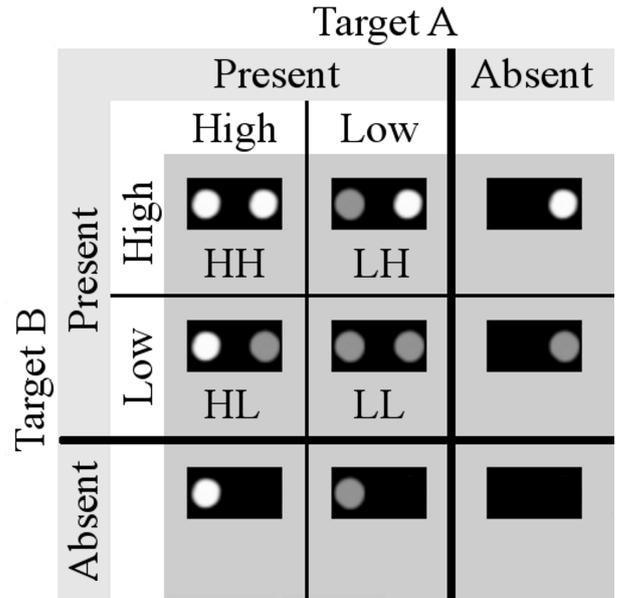


Figure 2: The Double Factorial Design. One level of manipulation is the presence or absence of a target within a channel. The other level is the salience of the present target: High salience (H) or Low salience (L). Bright and dim dots are used to illustrate a possible salience manipulation.

requiring the participant to only respond when dots are presented to both visual fields forces an exhaustive stopping rule. However, participants may exhaustively process all available information even if not all of the information is necessary for the task (e.g., Johnson et al., 2010). The AND and OR stopping rules are easily generalized to systems with more sub-processes, but with more sub-processes there are more possibilities, such as requiring some subset of sub-processes to stop.

Townsend and colleagues developed a set of tools for overcoming these aforementioned difficulties in classifying mental processes: Systems Factorial Technology (SFT). SFT consists of precise definitions of the properties of mental processes, including architecture and stopping rule as described above, as well as workload capacity and stochastic dependence. Workload capacity refers to how the speed of each individual sub-process changes as the number of sub-processes increases. Stochastic dependence is a measure of how the distributions of completion times of each of the sub-processes are related. For a more in depth development of these definitions, see (Townsend, 1972, 1974; Townsend and Ashby, 1983).

SFT includes an experimental paradigm, motivated by these theoretical concepts, for testing systems with two sub-processes of interest (Townsend and Nozawa, 1995).¹ This paradigm, the Double Factorial Paradigm (DFP), requires the manipulation of two factors, the activation of

¹In general SFT can apply to any number of sub-processes, although the majority of the work has focused on two sub-processes.

a sub-process and the speed of the active sub-process(es). These are referred to as the presence/absence and salience manipulations respectively. The factorial manipulations are summarized in Figure 2, along with stimuli similar to those used in the dot perception experiment mentioned earlier (Eidels et al., 2010) as an example of these manipulations.² Bright dots are perceived more quickly, and thus represent the high salience condition compared with the dimmer, more grey dots.

The SIC is calculated based on the effects of the salience manipulation in trials with both targets are present. In each of these four conditions, high-high, high-low, low-high, and low-low, the survivor function of the response times is approximated, which enters into computations of the SIC. The survivor function, $S(t)$, of the response times is the probability that the participant has not yet responded, so at time 0, $S(t) = 1$ and $S(t) \rightarrow 0$ as $t \rightarrow \infty$. It is the complement of the more familiar cumulative distribution function, $S(t) = 1 - F(t)$.³

Subscripts are used to denote the salience level of each process, so $S_{HL}(t)$ is the probability that the participant has not yet responded by time t when the first target is high salience (H) and the second target is low salience (L). Using this notation, the SIC is given by the following equation.

$$\begin{aligned} \text{SIC}(t) &= [S_{LL}(t) - S_{LH}(t)] - [S_{HL}(t) - S_{HH}(t)] \\ &= [F_{LH}(t) - F_{LL}(t)] - [F_{HH}(t) - F_{HL}(t)] \end{aligned}$$

To make conclusions based on the SIC, it is necessary that the salience manipulations selectively influence the sub-processes (Townsend and Nozawa, 1995). Informally, an experimental manipulation selectively influences a sub-process if that manipulation affects that sub-process and none of other sub-processes being measured. For a discussion of more formal definitions of selective influence, see (Dzhafarov, 2003; Ashby and Townsend, 1980). Under certain constraints on the distributions, direct tests for selective influence exist (Kujala and Dzhafarov, 2008), but in general we can only test that a set of empirical consequences of selective influence are not violated in the data. The most common of these tests is a test of the stochastic ordering of the survivor functions across conditions. This is summarized in the following inequalities, where $A < B$ is interpreted to mean that the random function A is stochastically less than the random function B . Each of these orderings can be tested for statistical significance using any test of distribution ordering, such as a one-sided Kolmogorov-Smirnov (K-S) test.

$$\begin{aligned} \tilde{S}_{HH}(t) &< \tilde{S}_{HL}(t) < \tilde{S}_{LL}(t) \\ \tilde{S}_{HH}(t) &< \tilde{S}_{LH}(t) < \tilde{S}_{LL}(t) \end{aligned}$$

²These were not the actual stimuli used, and are for illustration only. Eidels et al. (2010) used much dimmer dots for both conditions.

³The theory underlying SIC analysis does not focus on accuracy, so in practice the SIC is calculated from only correct responses.

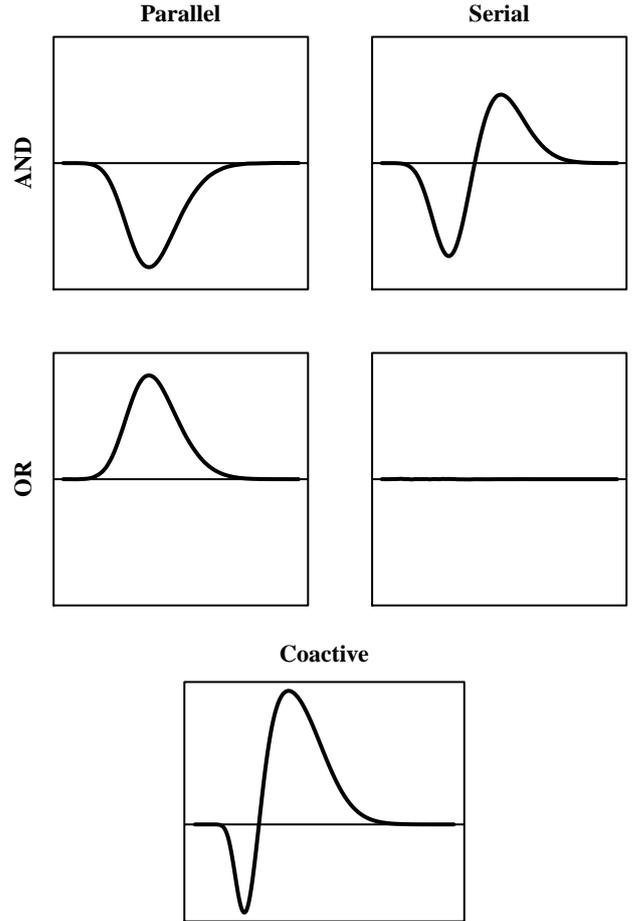


Figure 3: SIC predictions for the four canonical models plus the coactive model.

Under the assumption of selective influence of the salience manipulation and independence, each architecture–stopping rule combination predicts a distinct form of the SIC, depicted in Figure 3. The parallel-AND model predicts an entirely negative SIC, shown in the top left, whereas the parallel-OR model predicts an entirely positive SIC, shown in the bottom left. The serial-AND predicts both positive and negative parts of the SIC, as depicted in the top right, while the serial-OR model predicts the flat SIC depicted in the bottom right. Coactive models predict a small negative component of the SIC followed by a relatively larger positive part, with more positive area under the curve than negative.

These SIC predictions are extremely general as they make no assumptions about the distributions of the finishing times of each sub-process. Furthermore, the inclusion of a random variable representing the amount of time taken up by other processes involved in a response (base time) has no effect on these predictions (Townsend and Nozawa, 1995). Each of the four canonical architecture–stopping rule combinations predict consistent, distinct SIC shapes regardless of the finishing time distributions on

each of the channels. Although there is not a proof that the coactive SIC form in Figure 3 holds for all coactive models, it is the SIC predicted by all of the coactive models we have tested.

One limitation to SIC analysis is that until now there has been no measure of the statistical significance of a measured SIC. For example, if the empirical SIC is entirely positive except for some small negative deviation, it has been impossible to determine if that negative deviation is enough to falsify a parallel-OR model of performance, or for that matter if the positive deviation is enough to falsify a serial model or a parallel-AND model. The delay in the development of such a statistic thus far is due in part to the constraint that SIC analysis remain non-parametric. In this paper we develop a series of non-parametric tests to address this shortcoming. We begin by stating the necessary definitions along with the main theoretical results, then use those results to establish statistical tests.

1. Theoretical Approach

To find the appropriate non-parametric statistics, we generalize the basic findings related to empirical cumulative distribution functions (ECDF, Parzen (1962); Doob (1949)), using an approach similar to Maris and Maris (2003).

Definition 1. Let n be the number of trials for which a response time is collected and $I_{T_i \leq t}$ be the indicator variable such that

$$I_{\{T_i \leq t\}} = \begin{cases} 1 & T_i \leq t \\ 0 & T_i > t. \end{cases}$$

Then the ECDF is

$$\tilde{F}^{(n)}(t) = \frac{\sum_{i=1}^n I_{\{T_i \leq t\}}}{n}.$$

The statistical properties of the ECDF are well established. First, the ECDF is an unbiased and consistent estimator of the true CDF (Bickel and Spieler, 1977). Second, the asymptotic distribution of the error of the ECDF is known: Assuming the number of response times collected is not random, then the distribution of the supremum difference between the true CDF and the ECDF over the support of the CDF is equal in distribution to a Brownian Bridge (Theorem 1 below). A Brownian bridge is a special case of Brownian motion in which both the beginning and the end are fixed. Brownian motion and Brownian bridges are defined more formally in Definitions 2 and 3 respectively.

Definition 2. A random process $\{X(t), t \geq 0\}$ is a Brownian motion process with diffusion coefficient $\sigma^2 > 0$ if the following properties hold.

- $X(0) = 0$.
- $X(t)$ has independent, normally distributed increments with $E[X(t)] = 0$ and $E[X(t)X(s)] = \sigma^2 \min(t, s)$.

- $X(t)$ has continuous sample paths with probability 1.

Definition 3. A random process $\{\mathcal{B}(t), 0 \leq t \leq 1\}$ is a Brownian bridge process with diffusion coefficient $\sigma^2 > 0$ if $\{\mathcal{B}(t), t \geq 0\}$ is a Brownian motion process with diffusion coefficient σ^2 and $\mathcal{B}(0) = \mathcal{B}(1) = 0$.

We use $\mathcal{B}(t)$ to denote a standard Brownian bridge, in which the diffusion coefficient is 1 and $\mathcal{B}^{\sigma^2}(t)$ if $\sigma^2 \neq 1$.

The first theorem is a standard result, due originally to Kolmogorov (1933). For a relatively straightforward proof, see Parzen (1962, pp. 99–102) or Doob (1949). For a more in depth treatment see Billingsley (1999, p. 149).

Theorem 1. Let $\tilde{F}^{(n)}(t)$ be an ECDF calculated from n trials by the definition above. Then,

$$\sqrt{n} \sup_t \left| \tilde{F}^{(n)}(t) - F(t) \right| \xrightarrow{d} \sup_t |\mathcal{B}(t)| \text{ as } n \rightarrow \infty.$$

Equivalently,

$$\sup_t \left| \tilde{F}^{(n)}(t) - F(t) \right| \xrightarrow{d} \sup_t \frac{1}{\sqrt{n}} |\mathcal{B}(t)| \text{ as } n \rightarrow \infty,$$

or,

$$\sup_t \left| \tilde{F}^{(n)}(t) - F(t) \right| \xrightarrow{d} \sup_t |\mathcal{B}^{1/n}(t)| \text{ as } n \rightarrow \infty.$$

Furthermore,

$$\sup_t \left(\tilde{F}^{(n)}(t) - F(t) \right) \xrightarrow{d} \sup_t \mathcal{B}^{1/n}(t) \text{ as } n \rightarrow \infty$$

and

$$\inf_t \left(\tilde{F}^{(n)}(t) - F(t) \right) \xrightarrow{d} \inf_t \mathcal{B}^{1/n}(t) \text{ as } n \rightarrow \infty.$$

The empirical survivor interaction contrast (ESIC) is related to the SIC in a manner similar to the relationship between the ECDF and the CDF. Suppose k is the number of observations used to estimate $F_{LH}(t)$, l is the number of observations used to estimate $F_{LL}(t)$, m is the number of observations used to estimate $F_{HH}(t)$, and n is the number of observations used to estimate $F_{HL}(t)$. Then,

$$\widetilde{\text{SIC}}_{k,l,m,n}(t) = \left[\tilde{F}_{LH}^{(k)}(t) - \tilde{F}_{LL}^{(l)}(t) \right] - \left[\tilde{F}_{HH}^{(m)}(t) - \tilde{F}_{HL}^{(n)}(t) \right].$$

We estimate the ESIC using only the correct response times, since those are the distributions of interest. The ESIC is a consistent and unbiased estimator of the true SIC. Also, the supremum and infimum of the error in estimating the SIC converge in distribution to the supremum and infimum respectively of a Brownian bridge as the number of observations increases. These results are proven in the next three theorems.

Theorem 2. $\widetilde{\text{SIC}}(t)$ is an unbiased estimator of $\text{SIC}(t)$.

Proof. This follows trivially from the fact that the ECDF is unbiased and the linearity of the expectation operator.

$$\begin{aligned}
& \mathbb{E} [\widetilde{\text{SIC}}_{k,l,m,n}(t)] \\
&= \mathbb{E} \left[\left[\tilde{\mathbb{F}}_{\text{LH}}^{(k)}(t) - \tilde{\mathbb{F}}_{\text{LL}}^{(l)}(t) \right] - \left[\tilde{\mathbb{F}}_{\text{HH}}^{(m)}(t) - \tilde{\mathbb{F}}_{\text{HL}}^{(n)}(t) \right] \right] \\
&= \left[\mathbb{E} \left[\tilde{\mathbb{F}}_{\text{LH}}^{(k)}(t) \right] - \mathbb{E} \left[\tilde{\mathbb{F}}_{\text{LL}}^{(l)}(t) \right] \right] \\
&\quad - \left[\mathbb{E} \left[\tilde{\mathbb{F}}_{\text{HH}}^{(m)}(t) \right] - \mathbb{E} \left[\tilde{\mathbb{F}}_{\text{HL}}^{(n)}(t) \right] \right] \\
&= [\mathbb{F}_{\text{LH}}(t) - \mathbb{F}_{\text{LL}}(t)] - [\mathbb{F}_{\text{HH}}(t) - \mathbb{F}_{\text{HL}}(t)] \\
&= \text{SIC}(t)
\end{aligned}$$

□

A useful quantity in the following two theorems is a combination of the number of observations from each of the distributions that make up the ESIC,

$$N = \frac{1}{1/k + 1/l + 1/m + 1/n}.$$

Theorem 3. $\widetilde{\text{SIC}}_{k,l,m,n}(t)$ converges almost uniformly to $\text{SIC}(t)$ as $N \rightarrow \infty$ and is therefore a consistent estimator of $\text{SIC}(t)$.

Proof. $N \rightarrow \infty$ implies that each of $k, l, m, n \rightarrow \infty$ since if, for example k did not, then $N \rightarrow 1/1/k = k < \infty$.

Let $\frac{\epsilon}{4} > 0$.

By the Glivenko-Cantelli Theorem (Billingsley, 1995, pg. 269), each of the ECDF's converge almost uniformly in \mathbb{R} to the true CDF as the number of observations used increases. Thus, for each condition $\text{XY} \in \{\text{LH}, \text{LL}, \text{HH}, \text{HL}\}$,

$$\lim_{h \rightarrow \infty} \Pr \left\{ \left| \tilde{\mathbb{F}}_{\text{XY}}^{(h)}(t) - \mathbb{F}(t) \right| < \frac{\epsilon}{4} \right\} = 1$$

The ECDF's are independent, so

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \Pr \left\{ \left[\left| \tilde{\mathbb{F}}_{\text{LH}}^{(k)}(t) - \mathbb{F}_{\text{LH}}(t) \right| < \frac{\epsilon}{4} \right] \right. \\
&\quad \cap \left[\left| \tilde{\mathbb{F}}_{\text{LL}}^{(l)}(t) - \mathbb{F}_{\text{LL}}(t) \right| < \frac{\epsilon}{4} \right] \\
&\quad \cap \left[\left| \tilde{\mathbb{F}}_{\text{HH}}^{(m)}(t) - \mathbb{F}_{\text{HH}}(t) \right| < \frac{\epsilon}{4} \right] \\
&\quad \left. \cap \left[\left| \tilde{\mathbb{F}}_{\text{HL}}^{(n)}(t) - \mathbb{F}_{\text{HL}}(t) \right| < \frac{\epsilon}{4} \right] \right\} = 1.
\end{aligned}$$

This implies that the probability that the sum of the absolute differences is less than ϵ is 1.

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \Pr \left\{ \left| \tilde{\mathbb{F}}_{\text{LH}}^{(k)}(t) - \mathbb{F}_{\text{LH}}(t) \right| \right. \\
&\quad + \left| \tilde{\mathbb{F}}_{\text{LL}}^{(l)}(t) - \mathbb{F}_{\text{LL}}(t) \right| \\
&\quad + \left| \tilde{\mathbb{F}}_{\text{HH}}^{(m)}(t) - \mathbb{F}_{\text{HH}}(t) \right| \\
&\quad \left. + \left| \tilde{\mathbb{F}}_{\text{HL}}^{(n)}(t) - \mathbb{F}_{\text{HL}}(t) \right| < \epsilon \right\} = 1
\end{aligned}$$

Since the absolute value of the sum or difference is less than the sum of the absolute values,

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \Pr \left\{ \left| \tilde{\mathbb{F}}_{\text{LH}}^{(k)}(t) - \mathbb{F}_{\text{LH}}(t) \right. \right. \\
&\quad - \tilde{\mathbb{F}}_{\text{LL}}^{(l)}(t) + \mathbb{F}_{\text{LL}}(t) \\
&\quad - \tilde{\mathbb{F}}_{\text{HH}}^{(m)}(t) + \mathbb{F}_{\text{HH}}(t) \\
&\quad \left. \left. + \tilde{\mathbb{F}}_{\text{HL}}^{(n)}(t) - \mathbb{F}_{\text{HL}}(t) \right| < \epsilon \right\} = 1.
\end{aligned}$$

Therefore, by rearranging the terms,

$$\lim_{N \rightarrow \infty} \Pr \left\{ \left| \widetilde{\text{SIC}}_N(t) - \text{SIC}(t) \right| < \epsilon \right\} = 1$$

□

Theorem 4. The supremum, infimum and supremum of the absolute value of $[\widetilde{\text{SIC}}_{k,l,m,n}(t) - \text{SIC}(t)]$ respectively converge in distribution to the supremum, infimum and supremum of the absolute value of Brownian bridge process with diffusion coefficient $1/N$ as $N \rightarrow \infty$.

Proof. By Theorem 1,

$$\begin{aligned}
& \sup_t \left[\tilde{\mathbb{F}}_{\text{LL}}(t) - \mathbb{F}_{\text{LL}}(t) \right] \xrightarrow{d} \sup_t \mathcal{B}^{1/k}(t) \\
& \sup_t \left[\tilde{\mathbb{F}}_{\text{LH}}(t) - \mathbb{F}_{\text{LH}}(t) \right] \xrightarrow{d} \sup_t \mathcal{B}^{1/l}(t) \\
& \sup_t \left[\tilde{\mathbb{F}}_{\text{HL}}(t) - \mathbb{F}_{\text{HL}}(t) \right] \xrightarrow{d} \sup_t \mathcal{B}^{1/m}(t) \\
& \sup_t \left[\tilde{\mathbb{F}}_{\text{HH}}(t) - \mathbb{F}_{\text{HH}}(t) \right] \xrightarrow{d} \sup_t \mathcal{B}^{1/n}(t).
\end{aligned}$$

The sum, or difference, of independent Brownian bridge processes is again Brownian bridge processes, and just as variance adds with Normal random variables, diffusion coefficients add with Brownian motion processes.⁴ Hence,

$$\begin{aligned}
& \sup_t l \left[\widetilde{\text{SIC}}_{k,l,m,n}(t) - \text{SIC}(t) \right] \xrightarrow{d} \sup_t \left[\mathcal{B}^{1/k+1/l+1/m+1/n}(t) \right] \\
& \quad = \sup_t \left[\frac{1}{\sqrt{N}} \mathcal{B}(t) \right].
\end{aligned}$$

The same argument holds for the infimum and the supremum of the absolute value of the difference. □

The final theorem that we need is the distribution over the maximum and minimum values of a Brownian Bridge. This is a standard result, due originally to Kolmogorov (1933), so the proof is omitted. For the interested reader, the derivation of these distributions are available in Billingsley (1999, Section 9).

⁴Briefly this follows the facts that 1) the sum of continuous functions is continuous and 2) each of the increments in each summand is normally distributed, so each of the increments in the sum is normally distributed. To see that the variance adds, $\mathbb{E}[(X(t)+Y(t))(X(s)+Y(s))] = \mathbb{E}[X(t)X(s)+X(t)Y(s)+Y(t)X(s)+Y(t)Y(s)] = \mathbb{E}[X(t)X(s)] + \mathbb{E}[Y(t)Y(s)]$.

Theorem 5. Let D be the maximum deviation from zero of a Brownian Bridge, D^+ be the largest positive value of a Brownian Bridge, and D^- be the magnitude of the (minimum) of a Brownian Bridge,

$$\begin{aligned} D &= \max_{0 \leq t \leq 1} |\mathcal{B}(t)| \\ D^+ &= \max_{0 \leq t \leq 1} \mathcal{B}(t) \\ D^- &= -\min_{0 \leq t \leq 1} \mathcal{B}(t). \end{aligned}$$

Then the distributions of D, D^+ , and D^- are as follows,⁵

$$\begin{aligned} S_D(x) &= \Pr\{D \geq x\} = 2 \sum_{j=1}^{\infty} (-1)^{j+1} e^{-2j^2 x^2} \\ S_{D^+}(x) &= \Pr\{D^+ \geq x\} = e^{-2x^2} \\ S_{D^-}(x) &= P\{D^- \geq x\} = e^{-2x^2}. \end{aligned}$$

1.1. Hypothesis Testing

Based on Theorem 4, testing for significant positive or negative values of the ESIC is the same as testing for the difference between two ECDF's using the K-S test. The null hypothesis is that the true SIC does not deviate from zero, which implies that the ESIC converges in distribution to a Brownian Bridge. In particular, suppose N is as above and the empirical deviations are denoted

$$\begin{aligned} \hat{D}_{k,l,m,n} &= \max_{0 \leq t \leq 1} |\widetilde{\text{SIC}}_{k,l,m,n}(t) - \text{SIC}(t)| \\ \hat{D}_{k,l,m,n}^+ &= \max_{0 \leq t \leq 1} \widetilde{\text{SIC}}_{k,l,m,n}(t) - \text{SIC}(t) \\ \hat{D}_{k,l,m,n}^- &= -\min_{0 \leq t \leq 1} \widetilde{\text{SIC}}_{k,l,m,n}(t) - \text{SIC}(t). \end{aligned}$$

Then by Theorem 4 and Theorem 5, under the null hypothesis the limit distributions of the empirical deviations are given by,

$$\lim_{N \rightarrow \infty} \Pr\{\sqrt{N}D_N \geq x\} = 2 \sum_{j=1}^{\infty} (-1)^{j+1} e^{-2j^2 x^2} \quad (1)$$

$$\lim_{N \rightarrow \infty} \Pr\{\sqrt{N}D^+ \geq x\} = \Pr\{\sqrt{N}D^- \geq x\} = e^{-2x^2}. \quad (2)$$

These equations allow us to calculate the p -values and critical values for hypothesis testing on \hat{D}^+ and \hat{D}^- .

$$\begin{aligned} p_+ &= \exp\left[-2N \left(\hat{D}^+\right)^2\right] \\ p_- &= \exp\left[-2N \left(\hat{D}^-\right)^2\right] \\ \hat{D}^+_{crit} &= \hat{D}^-_{crit} = \sqrt{\frac{\log \alpha}{-2N}} \end{aligned}$$

The distribution of D (Equation 1) contains an infinite

⁵Note that the distributions of $S_{D^+}(x)$ and $S_{D^-}(x)$ are Weibull distributions with scale= $1/\sqrt{2}$ and shape= 2 .

α	.1	.05	.01
Critical value	$1.22/\sqrt{N}$	$1.36/\sqrt{N}$	$1.63/\sqrt{N}$

Table 1: Critical values for D statistic.

Model	\hat{D}	\hat{D}^+	\hat{D}^-	MIC
Serial-OR	\emptyset	\emptyset	\emptyset	\emptyset
Serial-AND	✓	✓	✓	\emptyset
Parallel-OR	✓	✓	\emptyset	✓
Parallel-AND	✓	\emptyset	✓	✓
Coactive	✓	✓	✓	✓
✓: Reject null hypothesis \emptyset : Fail to reject null hypothesis				

Table 2: Expected results of testing deviations for significance given each of the architecture-stopping rule combinations. The null-hypothesis is rejected when the appropriate p -value, e.g. $S_D(\hat{D}/\sqrt{N})$, is less than the chosen α . For clarity, the subscript of k, l, m, n was left off of the \hat{D} s.

sum, so we do not give the corresponding equations. In most cases, it is sufficient to use values from a table, such as Birnbaum (1952). For reference, approximate critical values are given in Table 1. There are also algorithms available for finding very good approximations to the distribution Marsaglia et al. (2003).

Depending on the information processing model we are testing, the different empirical deviations will be more or less informative. The expected results are summarized in Table 2. As an example, suppose we are interested in determining if an observed ESIC is representative of parallel-OR processing. According to Table 2, we should be able to reject the null-hypothesis that the overall deviation (\hat{D}) and the positive deviation (\hat{D}^+) are due to chance, but not the null hypothesis that the negative deviation (\hat{D}^-) is due to chance. In fact, we gain more information from these tests as well. If we find that there is a significant positive deviation, then by rejecting the null hypothesis, we also rule out both serial-OR and the parallel-AND classes of models.

A quick glance at Table 2 shows that the D statistics are not enough to distinguish among the all of the models. In particular, both coactive and serial-AND processing models predict significant positive and negative deviations. Fortunately, these models can be easily discerned using a test for an interaction between the salience manipulations on each channel. According to Townsend and Nozawa (1995), serial models predict no significant interaction, while a coactive model predicts a significant interaction. We define the mean interaction contrast, using M to denote mean as follows,

$$\text{MIC} = [M_{LL} - M_{LH}] - [M_{HL} - M_{HH}]$$

From the fact that for any positive random variable X , $\int_0^{\infty} \Pr[X > t] dt = E[X]$, we see that $\text{MIC} = \int_0^{\infty} \text{SIC}(t) dt$.

Thus, we see that a non-zero MIC is equivalent to a difference between the positive and negative areas of the SIC.

One possible test for an interaction in the means is an ANOVA, which assumes that the response times are normally distributed. Although the ANOVA is thought to be robust against violations of the normality assumption (Lindman, 1974), we again prefer the non-parametric approach. Following the suggestion of Sawilowsky (1990), we use the adjusted rank transform (ART) test. In the analyses below, both the ANOVA and ART results are included.

2. Simulation

We now turn to simulated data to empirically validate these results, as well as gain information on the power of the new test statistics. To generate these data, we simulated independent stochastic accumulators for each channel, similar to independent versions of the dynamic systems described in Townsend and Wenger (2004). At each time step Δt , the activation in a channel was updated according to the following equation,

$$x(t + \Delta t) = x + ax + u + \sqrt{\Delta t}N(0, 1). \quad (3)$$

For these simulations, Δt was set to .01.

In each round of the simulation, the input u_l for the low salience condition was drawn from a uniform distribution separately for each channel: For the parallel and serial models $u_l \sim \text{Unif}(1.5, 2.5)$, and the coactive model $u_l \sim \text{Unif}(2, 3)$. The high salience input, u_h was set to 1.5 times the low salience input. The accumulator had a small leak, $a = -.05$ so that the system would remain stable in the absence of a threshold.⁶

For each of the parallel/serial-OR/AND models, the channels completed when the accumulated activation reaches a threshold θ . Completion times were determined based on the appropriate model, e.g., the completion time for the parallel-OR model was the minimum completion time of the two channels. The coactive model finished when the summed values of the accumulators reached threshold. The threshold was set to $\theta = 5$ for the serial and parallel models and $\theta = 10$ for the coactive model.

These parameters were chosen to generally satisfy two constraints. First, we chose values that produced reasonable response time distributions (when scaled by a factor of 100) for simple perceptual tasks. Second, we attempted to minimize the number of times effective selective influence failed. Conclusions using the SIC are only valid if they are calculated from data that pass this test, so we only used simulated data that passed those tests.

We tested each of the distribution orderings implied by selective influence using the two sample K-S test. Selective influence implies $S_{HH} \leq S_{HL}, S_{LH}$, so we required that

Model	D	D^+	D^-	ART	ANOVA
Serial-OR	0.058	0.050	0.049	0.255	0.089
Serial-AND	0.923	0.405	0.948	0.062	0.035
Parallel-OR	0.937	0.969	0.002	0.988	0.988
Parallel-AND	0.907	0.000	0.951	0.539	0.710
Coactive	0.999	0.999	0.884	0.888	0.897

Table 3: Simulated probabilities of finding significant values under the null hypothesis that the deviation is due to chance with $\alpha = 0.05$ for each test individually, e.g. $\Pr\{S_D(\hat{D}_{k,l,m,n}/\sqrt{N}) > .05\}$. A Type I error occurs when the deviation should be chance according to Table 2 but is declared significant. The power of the test is the probability of finding a significant value when the deviation should be declared significant according to Table 2.

$\max(S_{HL} - S_{HH})$ and $\max(S_{LH} - S_{HH})$ were significant and $\max(S_{HH} - S_{HL})$ and $\max(S_{HH} - S_{LH})$ were not. Similarly, for selective influence to hold $S_{HL}, S_{LH} \leq S_{LL}$, so we required that $\max(S_{LL} - S_{HL})$ and $\max(S_{LL} - S_{LH})$ were significant and $\max(S_{HL} - S_{LL})$ and $\max(S_{LH} - S_{LL})$ were not.

We ran the simulation until we had 2000 rounds that had passed the test for selective influence, each consisting of 200 trials per condition. The noise was re-sampled for each trial and the input was varied randomly across rounds. The other parameters (e.g. stabilizing constant, threshold, ratio of high activation to low activation) remained fixed.

Table 3 shows the results of using the new hypothesis test on the simulated data. Each value is the proportion of times out of the 2000 simulations that the null hypothesis of random deviation was violated. In those cases where we expect to reject the null hypothesis, these values represent a Type I error rate. In those other cases, in which we expect to confirm the null hypothesis, these values represent the power. The results, particularly those representing power, are dependent on the parameters chosen. The most influential parameter is the ratio of high to low activation inputs. Higher ratios lead to more power and smaller ratios lead to less power.

The true SIC of a serial-OR process is flat, so any deviation from zero will be due to noise in the estimate. Since the distribution of this noise is approximated by a Brownian bridge in the limit, it is no surprise that the simulated serial-OR data Type I error rates were quite close to the chosen α level. Significant positive deviations of parallel-AND models and significant negative deviations of parallel-OR models were rare, far less frequent than the .05 associated with the chosen α -level. This is because each of these models predict SIC's that are in the opposite direction, making the null hypothesis of zero deviation more conservative.

Simulations of all of the models other than the serial-OR led to significant absolute deviations from zero (D) more than 90% of the time. Nearly all of the parallel-OR and coactive simulations had significant positive deviations

⁶Natural systems are often assumed to be stable, so the leak was added but it had no noticeable effect on the results.

	Serial OR	Serial AND	Parallel OR	Parallel AND	Coactive
Serial OR	0.901	0.000	0.050	0.049	0.000
Serial AND	0.035	0.350	0.018	0.561	0.037
Parallel OR	0.031	0.000	0.968	0.001	0.000
Parallel AND	0.049	0.000	0.000	0.951	0.000
Coactive	0.000	0.110	0.116	0.001	0.773

Table 4: Model confusions using the joint hypothesis given in Table 2.

(D^+) while the majority of the serial-AND, parallel-AND, and coactive model simulations had significant negative deviations, indicating high power for the statistic in each of these cases. The one case in which the test had relatively low power was the serial-AND model’s positive deviation.

There was some α inflation in the mean interaction tests, particularly the ART. The serial-OR model MIC was significant 26% of the time according to ART and 8% of the time using the ANOVA. Based on the chosen α it should have been approximately 5%. The tests were better on the serial-AND MIC, with 6% and 4% of the rounds significant using the ART and ANOVA respectively. The tests did have relatively good power on the parallel-OR and coactive models, but not as good for the parallel-AND model. Both ART and ANOVA detected significant MIC’s on the parallel-OR model in 99% of the rounds. The ART detected significant MIC’s for 88% of the coactive rounds and the ANOVA detected significance in 90%. For the parallel-AND model, the ART only detected it in 54%, while the ANOVA fared better with 71%.

Although both the ART and the ANOVA show some deficits with these data, the major shortcomings are not on the most important comparisons. The only two models that can not be distinguished using D^+ and D^- are the coactive and serial-AND models. Both the ART and the ANOVA had α levels near .05 on serial-AND and high power on the coactive models.

The main focus of these tests is to distinguish models, not just to determine deviations from zero in isolation. Hence, we are interested in how well these test perform in combination to distinguish models. Table 4 shows model confusion probabilities between each of the model types based on the predicted D^+ , D^- and MIC values given in Table 2. The ART was used to test the MIC and it was only used to distinguish between coactive and serial-AND models. Although the ANOVA performed slightly better with these models, the ART is a non-parametric test and hence the results should be less dependent on the particular model we simulated. The confusions in Table 4 are as expected from the Type I and power given in Table 3. In particular, the most common error these tests make is

classifying serial-AND models as parallel-AND due to the relatively low percentage of significant D^+ values.

Just as in the simulations, there is a practical way to increase the power of these tests. As always, more data results in more power. From Theorem 4, it is clear that as the sample size increases, the diffusion coefficient decreases and thus shrinks the largest likely random deviations. Thus there is a smaller range of deviations that are plausibly due to chance. Alternatively, increasing the strength of the salience manipulations, similar to increasing the ratio of the high to low salience input, will increase the power. Townsend and Diederich (1996) showed that increasing the effect of the salience manipulation increases the predicted deviations from zero in the SIC. Hence, when focusing on those tests that have relatively lower power, it is important to collect as much data as is reasonable and ensure that the salience manipulation is as effective as possible.

3. Application

To demonstrate the use of this statistical testing procedure we apply it to data collected from a DFP experiment examining the combined visual processing of the right and left visual fields (Eidels et al., 2010). The task is the one described in the earlier example and summarized in Figure 2. Participants were shown a small grey dot on a black background on either the left side of the screen, the right right side, or both (the presence/absence manipulation). When dots were displayed, they were shown at one of two levels of brightness (the salience manipulation). Participants completed two versions of each task. In one version they responded “yes” when either dot was present and “no” otherwise (OR task). In the other version they responded “yes” only when both dots were present and “no” otherwise (AND task). For more details on the experimental setup and participants, see Eidels et al. (2010).

Figure 4 shows the ESIC for one participant who had particularly clear results on both the OR task and the AND task. In both versions of the task, the ESIC is entirely to one side of the axis, conforming to the prediction of a parallel model. The ESIC from the OR task, on

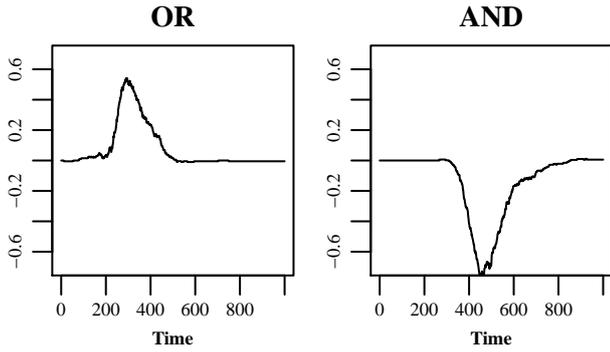


Figure 4: ESIC(t) for one participant (WY) on the OR and AND version of the dots task.

the left, is entirely positive as predicted by a parallel-OR model. The ESIC from the AND task, on the right, is entirely negative as predicted by a parallel-AND model. These results are reasonable given that serial processing of the left and right visual field seems unlikely. While a coactive model may be a reasonable model for this task a priori, neither SIC indicates such a process.

Given these interpretations of the data based on the visual attributes of the ESIC, we would now like to confirm that those interpretations are in fact statistically significant. Although each participant had 200 trials with each of the salience combinations, we are limited to correct response times. The ESIC for the OR task was estimated from survivor functions with 192, 198, 198 and 199 response times. Thus, $N = 49.2$ and the critical value of the test statistics \hat{D}^+ and \hat{D}^- at $\alpha = .01$ is 0.216. The maximum positive deviation from zero of this ESIC is $\hat{D}^+ = 0.54$, which is larger than the critical value. Hence, we reject the null hypothesis that the true SIC is flat and the positive deviation is due to chance. Referring to Table 2, we see that this result falsifies serial-OR and parallel-AND models of performance on this task. The maximum negative deviation is $\hat{D}^- = .001$, which is not significant (p -value = .99). This is evidence in favor of the parallel-OR model, since the coactive and serial-AND models predict significant negative deviation.

The ESIC for the AND task was estimated from survivor functions with 171, 165, 156 and 191 data points each, so $N = 42.5$ and the critical value for the test statistics at $\alpha = .01$ is 0.233. The maximum negative deviation of the ESIC is $\hat{D}^- = .763$, so again we can reject the null hypothesis that the true SIC is flat. According to Table 2 this implies that the serial-OR and parallel-OR models are not reasonable for these data. As evidence in favor of the parallel-AND model, the maximum positive deviation was only $\hat{D}^+ = .011$ and hence not significant (p -value = .99).

Table 5 shows the maximum deviations and corresponding p -values from the test statistic for each participant on both of the tasks. Nearly all of the participants follow the same pattern as the participant mentioned above (Partic-

ipant WY). The only exception is Participant BJ who's SIC was not significantly positive on the OR task. In this case we cannot rule out serial-OR as a model of processing on this task. None of the participants had both significant positive and negative parts of the ESIC, so coactive and serial-AND are not coactive models. Hence, the MIC is not necessary for classification and therefore not included. Based on these results it reasonable to conclude that people process information from the left and right visual fields in parallel, with a stopping rule corresponding to the task demands.

4. Discussion and Future Directions

In this paper we demonstrated the statistical properties of the empirical survivor interaction contrast (ESIC). In doing so, we have developed a significance test to aid in the use of the SIC as a tool for identifying mental processing characteristics. This test can be applied in its different forms to falsify any of the serial-OR, parallel-OR, and parallel-AND models of information processing. Based on both statistical theory and simulation, we have found this test to be an effective significance test. Furthermore, we have demonstrated its application to a simple data set and found it to be quite powerful in this case.

The statistical tests we developed herein are limited in the same way that all null-hypothesis significance tests are. We cannot use these tests to directly compare between models. Often we are seeking to know whether a serial or a parallel model is a more reasonable description of the data. An alternative approach, using the Bayesian statistical framework may offer a remedy for this limitation. Using the theorems about the ESIC contained in this paper, we are pursuing this alternative approach.

Despite this limitation, these new developments remove a major barrier to the use of the survivor interaction contrast. Although the survivor interaction contrast is the best available tool for determining the architecture and stopping characteristics of mental processes, the lack of a method for quantitative testing has hindered its acceptance. The theories presented in this paper fill this need by demonstrating the statistical properties of the ESIC and providing significance tests based on those properties.

Acknowledgements

This work was supported by NIH-NIMH MH 057717-07 and AFOSR FA9550-07-1-0078. We would like to thank Mario Fifić, Leslie Blaha, Lei Pei, and Nicholas Altieri and reviewers Philip Smith, Andrew Heathcote and E. J. Wagenmakers for their insight and commentary.

References

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

Participant	OR Task			AND Task		
	\hat{D}/\sqrt{N}	\hat{D}^+/\sqrt{N}	\hat{D}^-/\sqrt{N}	\hat{D}/\sqrt{N}	\hat{D}^+/\sqrt{N}	\hat{D}^-/\sqrt{N}
BJ	0.16	0.16	0.00	0.40***	0.01	0.40***
RS	0.52***	0.52***	0.03	0.31***	0.02	0.31***
JS	0.30***	0.30***	0.11	0.48***	0.00	0.48***
MB	0.48***	0.48***	0.01	0.44***	0.03	0.44***
RM	0.20*	0.20*	0.02	0.30***	0.00	0.30***
LB	0.36***	0.36***	0.03	0.60***	0.03	0.60***
JG	0.68***	0.68***	0.02	0.57***	0.00	0.57***
WY	0.54***	0.54***	0.01	0.76***	0.01	0.76***
AW	0.69***	0.69***	0.02	0.71***	0.00	0.71***

Table 5: Results of the SIC test statistics applied to the results of the two dot experiment of Eidels et al. (2010).

- Bickel, P. J. and Spieler, D. H. (1977). *Mathematical Statistics*. Holden-Day, San Francisco.
- Billingsley, P. (1995). *Probability and Measure*. Wiley, New York, third edition.
- Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley, New York, second edition.
- Birnbaum, Z. W. (1952). Numerical tabulation of the distribution of Kolmogorov’s statistic for finite sample size. *Journal of the American Statistical Association*, 47(259):425–441.
- Doob, J. L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *The Annals of Mathematical Statistics*, 20:393–403.
- Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrika*, 68:7–26.
- Eidels, A., Townsend, J. T., Hughes, H. C., and Perry, L. A. (2010). Complementary relationship between response times, response accuracy, and task requirements in a parallel processing system. Manuscript submitted for publication.
- Johnson, S. A., Blaha, L. M., Houpt, J. W., and Townsend, J. T. (2010). Systems factorial technology provides new insights on globallocal information processing in autism spectrum disorders. *Journal of Mathematical Psychology*, 54:53–72.
- Kolmogorov, A. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell’ Istituto Italiano degli Attuari*, 4:83–91.
- Kujala, J. V. and Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, 52:128–144.
- Lindman, H. R. (1974). *Analysis of variance in complex experimental designs*. Freeman, W. H. and Company, San Francisco.
- Maris, G. and Maris, E. (2003). Testing the race model inequality: a nonparametric approach. *Journal of Mathematical Psychology*, 47:507–514.
- Marsaglia, G., Tsang, W. W., and Wang, J. (2003). Evaluating Kolmogorov’s distribution. *Journal of Statistical Software*, 8:1–4.
- Parzen, E. (1962). *Stochastic Processes*. Holden-Day, San Francisco.
- Sawilowsky, S. S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research*, 60:91–126.
- 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., Dann, C., and McGuigan, K. (1989). Inferring mental process organization from reaction times. In Vickers, D. and Smith, P. L., editors, *Human Information Processing: Measures, mechanisms, and models*. Elsevier Science, North-Holland.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153:652–654.
- 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 Kantowitz, B. H., editor, *Human Information Processing: Tutorials in Performance and Cognition*, pages 133–168. Erlbaum Press, Hillsdale, NJ.
- 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. (1990). 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. and Ashby, F. G. (1983). *The Stochastic modeling of elementary psychological processes*. Cambridge University Press, Cambridge.
- Townsend, J. T. and Diederich, A. (1996). The relationship of variance to interaction contrast in parallel systems factorial technology. *British Journal of Mathematical Psychology*, 49:211–223.
- Townsend, J. T. and Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*, 39:321–360.
- Townsend, J. T. and 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.