Published in: Fifić, M. (2016) Simple Factorial Tweezers for detecting delicate serial and parallel processes. Mathematical Models of Perception and Cognition: Essays in Honor of James T. Townsend" (J. W. Houpt & L. M. Blaha, Eds). In Press. New York: *Psychology Press*.

Simple Factorial Tweezers for detecting delicate serial and parallel processes

Mario Fifić

Grand Valley State University

Over the last fifty years, work regarding the theoretical foundations governing the organization of mental processes has centered on the formal properties of various hypothesized mental networks. Such networks are defined in terms of their fundamental properties: processing order, stopping rule, and process dependency. Pivoting on the work of James Townsend and other essential contributors such as Richard Schweickert and Ehtibar Dzhafarov, these efforts resulted in the creation of the Systems Factorial Technology (SFT) - a suite of methodological tools for directly investigating the fundamental properties of cognitive operations. The SFT approach rests on rigorously tested mathematical tools for discriminating between serial and parallel processing, exhaustive and self-terminating stopping rules, and stochastic independence and dependence, as well as for discerning the capacity of an investigated system, all in a non-parametric (distribution-free) manner. The present study is focused on further refining recent advances in the SFT methodology and on the development of new tools for use with mental networks consisting of more than two processes. The present study also seeks to integrate these advances with the factorial tools developed to explore non-homogeneous mental networks, which may consist of both serial and parallel processes (so called serial/parallel networks).

One of the essential tasks of cognitive psychology is to learn how mental processes are organized in various cognitive tasks. Over the last several decades, cognitive psychologists have been trying to validate various cognitive models by asking questions such as: What is the order of the completion of these mental processes (e.g., serial or parallel)? Can a cognitive system be terminated when only a few or all processes has been completed (e.g., self-termination and exhaustiveness)? Do processes of interest depend on each other (e.g., process interdependency)? What is a system's processing capacity (e.g., limited, unlimited, or super capacity). These questions focus on the aspects of processing which are all referred to as the fundamental properties of mental processes.

There has been a constant development of methodologies to uncover the fundamental properties of mental processes. An important early contributor, F. C. Donders (1868) devised a subtraction method which measures processing time durations. In this method, two tasks are used with the only difference between the tasks being that in the second task, an additional process has been inserted. The duration of the *inserted* mental process is inferred by subtracting the times needed to complete two tasks.

With the rise of the modern cognitive psychology in the 1960's, a new approach was developed; the "additive factors method" explored the fundamental properties of mental processes such as serial and parallel processing order. Unlike the subtraction method, which worked via inserting processes, the method proposed by Sternberg (1966; 1969) was based on affecting the duration of processes within an unknown mental network (with the motto "Stretching processes rather than inserting them", after Schweickert, Fisher & Sung, 2012). Sternberg proposed that, by the virtue of selective influence, it should be possible to elongate the processing durations. The selective influence was (and is) considered as the critical conditional assumption of the methodology. This served as the experimental manipulation which would affect the duration of only one of the processes and leave the other processes in the network unaffected. In the additive factors method, a factor is defined as an experimental variable which affects the duration of a single process of interest. The advantage of the additive factors over the subtraction method is that researchers do not need to change the structure of the mental network under investigation.

In practice, the additive factors method employs an ANOVA to test for an interaction between factors. Serial and parallel processing are tested simply by observing the absence or presence of the interaction between experimental factors, respectively.

In his seminal work on distinguishing between serial and parallel processing in short-term memory (STM), Sternberg (1966; 1969) proposed a combination of his additive factors and Donders's subtraction methods. Some processes were stretched (e.g., imposing mask on a target) and some other processes were inserted (e.g., by varying the number of digits to be memorized across different conditions –also called workload). In his short-term memory studies, Sternberg concluded a strict serial processing of memorized units based on the following findings: (a) The interaction effects of the key experimental factors were not significant thus showing additivity. From the Sternberg's point of view, the additive effects meant that the effects of either adding, subtracting or stretching processes independently affected overall response time in a STM task. The additivity supports the hypothesis that the selective influence held; and (b) The reaction time (RT) showed the linearly increasing trend as a function of the number of memorized items (linear RT function of workload), that is RT increased linearly as a function of the workload. This finding supported that the scanning time per one item was constant over the memory load and consequently supported the hypothesis that the memorized items were scanned in a serial fashion. Sternberg believed it would be possible to discriminate between pure serial and parallel processing by analyzing the shape of the RT-workload function: the landmark signature of the serial system should be a linearly increasing RT-workload function. That is by adding more processes the system should slow down at a constant rate. In contrast, the parallel system should exhibit a flat RT-workload function, that is, by adding more processes that are analyzed in parallel, the system should not slowdown in processing as all processes occur simultaneously.

The theoretical breakthrough

In the original Sternberg STM paradigm, one of the key experimental manipulations is the number of memorized units (workload). By manipulating a workload, a researcher directly increases or decreases a size of the mental network that make a short-term memory store and consequently affects a number of conducted processes. For example, if a subject has to memorize 4 items then the search for the target item should include N=4 of target-to-item comparisons.

Essentially, the workload manipulation is based on the insertion method which has been criticized in the post-Dondersian era as having the potential to inviting confounding variables (e.g., Townsend, 1971; Townsend & Ashby, 1983). For example, by inserting a number of memorized units, thus adding more items to be memorized, one would also be affecting the capacity of short-term memory storage. If the system is of limited capacity then it is to be expected that stored items should receive part of a shared amount of resources. So as workload increases it is likely that a cognitive system with limited capacity would distribute less processing resources per memorized unit. One could argue that such capacity limitations would not prevent the serial processing system from exhibiting its landmark signature - the linearly increasing RT-workload function. Provided that each processing unit receives the same amount of shared limited capacity the RT-workload function should increase linearly as more items are processed in a strictly serial fashion.

However, the parallel system under the same capacity limitation would fail to leave the classical parallel processing signatures: the flat RT-workload function. The problem is that quite naturally parallel models whose channels become less efficient as workload increases can make predictions identical to those of serial models —this is the well-known *model mimicking dilemma*. The groundbreaking work to solving the model mimicking was laid out by the work of James Townsend and his colleagues (e.g., Townsend, 1969; 1971; 1972; Townsend & Ashby, 1983, Chapter 14).

Pure stretching method

The work that followed Sternberg's studies was focused on designing the new methodologies to explore the fundamental properties of mental processes. Built on the work of James Townsend and other essential contributors such as Richard Schweickert and Ehtibar Dzahfarov the efforts resulted in a creation of the systems factorial technology (SFT) – a suite of methodological tools aimed at discovering the fundamental properties of cognitive operations. The SFT approach rests on rigorously tested mathematical tools for discerning serial from parallel processing exhaustive from self-terminating processing, process (in)dependence and the capacity of the system under investigation. During the three decades following the cognitive revolution in 1960s, James Townsend worked on refining the non-parametric mathematical methods that constitute the SFT suite of methodologies.

The breakthrough in the work of James Townsend (Townsend & Ashby, 1983; Townsend & Nozawa, 1995) was the theoretical definition and application of the so-called *pure stretching method* approach (Schweickert, Fisher & Sung, 2012); In the pure stretching approach no processes are inserted, and the analysis of the underlying mental network is conducted on a fixed number of processes. The pure stretching approach avoids possible confounds due to the capacity issue in exploring serial parallel system's properties using the insertion method. As a result, the pure stretching method improves the model selection by reducing a possible mimicking between different models.

Single factor manipulation: Stretching one process

To explain the effect of stretching of a certain mental process over time, I will define an internal psychological function h that affects the speed of mental process X by either slowing down or speeding up. I will also define a binary valued variable $x=\{x_{low}, x_{high}\}$, such that this set is ordered $x_{low} < x_{high}$. If, take that x is a fixed factor, then it is operationalized as an external manipulation which exclusively affects only the process of interest X, providing previously mentioned the selective influence. The effect of the manipulation of x on a speed of a mental process t_X is determined by the $h(X;x)$. The following conditions should apply so that the stretching manipulation could be used in SFT approach:

(a) Selective influence holds. This means that each experimental manipulation always works to affect only one process of interest (Sternberg, 1966; 1967).

(b) Stochastic order holds. The stochastic order of the effect of *h* function on a mental process duration is preserved for different magnitudes of external manipulation. Assume an underlying process of interest *X* which duration is affected by the *h(x)*. Assume that the process *X* is selectively influenced by binary valued experimental factor *x* across different levels (low and high) then it should always hold that $E[t_X; x_{low}] \geq E[t_X; x_{high}]$. In other words, the manipulation of a factor *x* at a low level should always lead to an equal or a slower expected processing rate of the process X then when *x* was at the high level. Even stronger test for the stochastic order (stochastic dominance) is conducted by looking at the order of corresponding survivor functions of reaction times, thus replacing the expectation times with the survivor function: $S(t_X; x_{low}) \geq S(t_X; x_{high})$.

(c) Process independence between processes holds which means that the rate of processing of a single process (say *X*) does not depend on any other process in a mental network (say *Y* and *Z*).

In practice, the stretching manipulation via the function *h*, is usually achieved either by selective visual masking (Sternberg, 1969), stimulus brightness (Townsend & Nozawa, 1995) in the cases of external visual search, or by inter-item similarity in the case of an internal memory search or categorization (Townsend & Nozawa, 1995; Townsend & Fific 2004, Fific, Little & Nosofsky, 2010).

The stretching effect of process *X* is presented here as a first-order difference operator Δ applied on the *h* function of the mental process *X*, for the parameter *x:* $\Delta h(X;x)$ (see also Townsend & Thomas, 1994). In case that the function operates at the expected time value of the process X:

$$\Delta E[t_X; x] = E[t_X; x_{low}] - E[t_X; x_{high}] > 0 \qquad (1)$$

Or at the survivor function level:

$$\Delta S(t_X; x) = S(t_X; x_{low}) - S(t_X; x_{high}) > 0 \qquad \text{for all } t_X$$

In practice, the first-order difference effect on mean reaction times indicates whether the mental process of interest was affected by experimental manipulation.

Double factorial manipulation: Stretching two process

Here we consider a minimal size mental network made of two processes that allows for assessment all fundamental properties. To assess properties of mental networks a researcher orthogonally combines the manipulated variable levels (for more details see Anderson & Whitcomb, 2000). In SFT, the *h* function that operates on two variables (t_X, t_Y) orthogonally combines the binary values of each experimental factor (low, high). Thus, the following <u>second-order</u> difference operator is defined:

$$\begin{aligned}
&\Delta^2 h(t_X, t_Y; x, y) = \Delta(\Delta h(t_X, t_Y; x_{low}, x_{high}, y_{low}, y_{high})) = \\
&\Delta(h(t_X, t_Y; x_{low}, y_{low}, y_{high}) - h(t_X, t_Y; x_{high}, y_{low}, y_{high})) = \\
&\Delta h(t_X, t_Y; x_{low}, y_{low}, y_{high}) - \Delta h(t_X, t_Y; x_{high}, y_{low}, y_{high}) = \\
&h(t_X, t_Y; x_{low}, y_{low}) - h(t_X, t_Y; x_{low}, y_{high}) - (h(t_X, t_Y; x_{high}, y_{low}) - h(t_X, t_Y; x_{high}, y_{high}))
\end{aligned} \qquad (2)$$

This second-order difference $\Delta^2 h(X, Y; x, y)$ is also known as the interaction contrast used to measure between-factor ANOVA interactions. The substantial literature is devoted to the nature of factorial tests in research (Anderson & Whitcomb, 2000, Maxwell & Delaney, 1999). The second-order difference in this form results in the four conditions corresponding to the four *h* function in the Eq 2. One can observe that four conditions are obtained by the orthogonal combination of the levels of the manipulated values of the factors (low and high). I will abbreviate the four conditions in the following fashion: LL, LH, HL and HH correspondingly to the last line of Eq 2. For example, LH means that the first manipulated stretching factor affecting the first variable t_X was at the low level=L, and the second manipulated stretching factor affecting the second variable t_Y was at the high level=H.

The second-order difference of a stretching effect provides sufficient information to explore the fundamental properties of the two mental process. First, it includes the two first-order differences, each providing information whether each of the experimental manipulation affected the process of interest through the stretching manipulation. Second, the information about the type of interaction between the two factors, through the second-order difference, can be used to distinguish between the fundamental properties.

SFT statistical tests for 2-process mental networks (N=2): MIC and SIC

The main testing tools in the SFT approach are the two statistics applied on both expected value of response time RT (*E[RT]*) and on corresponding time survivor functions (*S[t]*) over time *t*.

The second order difference Eq 2 could be applied on expected response time values, and it is known as the mean interaction contrast (MIC). The MIC statistic calculates the interaction between the factors, similarly to an ANOVA (Sternberg, 1969; see also Schweickert, 1978; Schweickert & Townsend, 1989):

To calculate MIC for the two processes, the second-order difference is derived for two variables, (X,Y) each belonging to distinct process within an unknown mental network:

$$\begin{aligned}
&\Delta^2 E[t_X, t_Y; x, y] = \\
&E[t_X, t_Y; x_{low}, y_{low}) - E[t_X, t_Y; x_{low}, y_{high}) - (E[t_X, t_Y; x_{high}, y_{low}] - E[t_X, t_Y; x_{high}, y_{high}])
\end{aligned}$$

In the case of multiple processes one has to verify whether the first-order difference of the first variable (t_X) aggregated over the levels of the other variable shows the expected mean order difference between low and high levels or not:

$$\Delta E[t_X, t_Y; x, y] =$$
$$E[t_X, t_Y; x_{low}, \mathrm{y}] - E[t_X, t_Y; x_{high}, \mathrm{y}]$$

When stretching is in place, the expected response time expectation is that the average response of the process X takes longer when the level of stretching manipulation was low than high, across all levels of the y process manipulations [1]:

$$E[t_X, t_Y; x_{low}, \mathrm{y}] - E[t_X, t_Y; x_{high}, \mathrm{y}] > 0$$

The analogous check is done across the levels of *y* factor. The strong stochastic order (stochastic dominance) is verified by looking at the order of corresponding survivor functions, thus replacing the expectations with the survivor functions, and checking whether the difference between two marginal survivor functions satisfy the inequality for both across *x* and *y*:

$$S(t_X, t_Y; x_{low}, \mathrm{y}) - S(t_X, t_Y; x_{high}, \mathrm{y}) > 0 \quad \text{for all } t_X$$

An even stronger test that the stochastic ordering is preserved is to check that the high and low effect on stretching of probability density functions cross exactly once (or an odd number of times (Townsend & Nozawa, 1995; Yang, Fific & Townsend 2014; Schweickert, Giorgini & Dzhafarov, 2000; Schweickert, Fisher & Sung, 2012).

The above second-order difference (Eq 2) on mean RTs can be written in the more popular form as so-called the double-factorial test in SFT:

$$MIC = (RT_{LL} - RT_{LH}) - (RT_{HL} - RT_{HH}) = RT_{LL} - RT_{LH} - RT_{HL} + RT_{HH} \quad (3)$$

RT stands for mean reaction time and the left and right subscripts refer to the stretching levels of the first and the second mental process of interest correspondingly. For example, HL indicates a condition where the first factor (processing the first item) is at the high level and the second factor (processing of the second item) is at the low level. The resulting design could be denoted as 2 × 2 factorial design (as employed in ANOVA).

Completely analogous to deriving the mean interaction contrast Eq 3, one can compute the survivor interaction contrast (SIC). By replacing the mean RTs for each condition by the survivor function symbol, at each value of *t*, one computes:

$$SIC(t) = (S_{LL}(t) - S_{LH}(t)) - (S_{HL}(t) - S_{HH}(t)) = S_{LL}(t) - S_{LH}(t) - S_{HL}(t) + S_{HH}(t) \quad (4)$$

In practice when the double-factorial SFT test is used, it is necessary to check whether the four factorial conditions satisfy the following order at the mean reaction times:

$$M(RT_{LL}) \geq M(RT_{LH}),\ M(RT_{HL}) \geq M(RT_{HH})$$

And, the stronger test by inspecting the order of survivor function. It is important to keep in mind that, in order to apply SFT, the ordered survivor functions should not intersect when plotted (for statistical tests see Houpt & Townsend, 2010; Houpt, Blaha, McIntire, Havig, & Townsend, 2014; Yang, et al. 2014):

$$S_{LL}(t) \geq S_{LH}(t),\ S_{HL}(t) \geq S_{HH}(t), \quad \text{for all } t$$

There is a close relation with the results from the MIC because the value of the MIC is simply the integral of the SIC, for all values of *t.* The integral of the survivor function for a random variable yields the mean of that random variable. Because the survivor interaction contrast is simply a linear contrast of individual survivor functions, the integral of the SIC is a linear contrast of the means of the corresponding random variables (i.e., the MIC).

The showcase of the distinct diagnostic predictions of MIC and SIC patterns for several types of mental networks, made of two processes (*N*=2) are displayed in Figure 1A, the first column. The depicted mental networks are could be characterized as canonical networks as they possess only one processing order and one stopping rule. MIC value is presented in each signature's upper right corner.

[1] In the language of ANOVA, this would correspond to the finding the main effect of the first variable (e.g. Maxwell & Delaney, 1999).

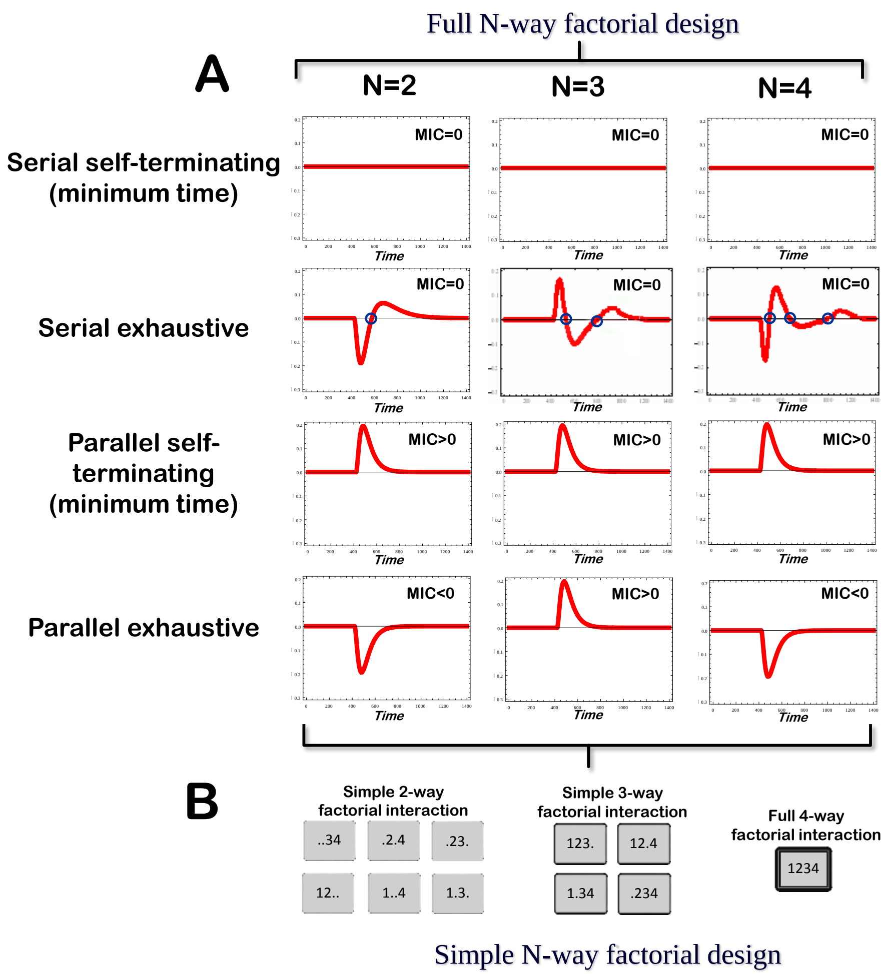

Figure 1: (A) The showcase of the distinct diagnostic predictions of MIC and SIC patterns for each distinct mental network (rows), of a size up to four processes (columns) for the full factorial research designs (top row). For the serial exhaustive mental network, the circles emphasize the intersection points of the SIC function and x-axis. (B) The bottom row indicate the type of simple interaction factorial test, derived from the 4-way full factorial design. Marginalized factors are denoted with a dot symbol. Each column indicate 2-way simple, 3-way simple and 4-way full factorial designs, and corresponding diagnostic predictions of MIC and SIC patterns. Note that all six simple 2-way simple interaction designs predict shape-equivalent SICs and MICs, for different mental networks. The same holds for the four 3-way simple interaction designs. Both (A) and (B) predicted MIC and SIC results are only relevant for the homogeneous mental networks (see the text for more details).

It is important to note that SIC provides more diagnostic power than MIC, as the SIC is a function of time, while MIC is the integrated value of that function over time . For example, certain classes of models cannot be distinguished using the MIC but can be successfully distinguished using the SIC [2].

Although the MIC test has less diagnostic power, MICs are more practical to use then SICs because one requires fewer trials to achieve stable estimates of the mean RT statistics than to get good estimate of the corresponding RT survivor function. So, there is a trade-off between practical applicability and the statistical power when using MIC and SIC.

From Equations 3 and 4 it is immediately clear that the SFT method, based on the pure stretching method, bypasses the confounding concerns that challenged the application of the Sternberg's insertion procedure in a short-

[2] Such a case would be the distinction between the parallel minimum time model on one side and the coactive model (or its close cousin - the parallel interactive model) on the other side. The MIC would predict positive values for all three model classes, but the SIC functions would differ in the small negative function values for early times (only the coactive and parallel interactive models predict early times t negative SIC values). Consideration of such cases is out of the scope of the current paper and more information is provided in other publications (Townsend & Nozawa, 1995).

term memory search task. By using SFT there is no need to insert processes to learn whether processing is serial or parallel, instead an experimental situation with a fixed number of process (N=2 in Figure 1A) is sufficient. This way, the SFT approach addresses the serial/parallel mimicking problem (mentioned above) due to the confounding of the workload manipulation and the network's capacity.

The SFT tools are primarily considered as being meta-theoretical tools (or a meta-theory). The SFT does not formally represent any specific model of cognitive operations. Instead SFT tools are primarily used for the operations of validation and/or falsification of cognitive models and also as exploratory tools to learning about new underlying cognitive operations.

The SFT models' validation/falsification power has been successfully utilized across different fields of cognitive psychology. The results challenged the standard expectations of some models to predict outcome of cognitive operations. For that purpose SFT has been used in the context of various cognitive tasks and domains: perceptual processes (e.g., Eidels, Townsend, & Pomerantz, 2008; Fific, Nosofsky, & Townsend, 2008; Johnson, Blaha, Houpt, & Townsend, 2010; Townsend & Nozawa, 1995; Yang, 2011; Yang, Chang, & Wu, 2013), visual and memory search tasks (e.g., Egeth & Dagenbach, 1991; Fific, Townsend, & Eidels, 2008; Sung, 2008; Townsend & Fific, 2004; Wenger & Townsend, 2001; 2006), face perception tasks (Fific & Townsend, 2010; Ingvalson & Wenger, 2005), and classification and categorization (e.g., Fific, Little, & Nosofsky, 2010; Little, Nosofsky, & Denton 2011; Little, Nosofsky, Donkin, & Denton, 2013). The SFT tools were recognized as potentially the most important and promising methodology in understanding cognitive processes (Greenwald, 2012).

Another use of the SFT tools is to define the minimal research design complexity criteria which is necessary to learn about cognitive operations of interest (Fific, 2014). In other words, the SFT can define the benchmark for the size of the research design which must be used to correctly recognize the network underlying cognitive operations. The question posed here is: How complex should a research design be in order to make valid inferences from the study about the fundamental mental properties (see Fific, 2014 for more details)? The answer to that question depends on the number of mental processes under investigation. So, to learn about a mental network of two processes (N=2), the SIC/MIC tests require at least two manipulated variables that are cross combined at the two process stretching levels (high and low). Then the number of experimental conditions is 2^2=4. To learn about the larger mental networks made of three processes, the number of experimental condition should be 2^3=8. In general, to test more complex mental networks made of N process, one has to employ at least 2^N experimental conditions.

SFT statistical tests for 2-process mental networks (N=2): The principle limitations

The development of the SFT tests for assaying networks of only two processes has helped in the validation of various cognitive models. However, the confinement of the SFT approaches to only N=2 two processes has limited the applicability of the SFT tools. Many cognitive tasks were originally designed to explore mental networks of larger sizes. For example, a short-term memory task which usually involves up modelling a search for up to 6 stored memorized unit, thus the proposed number of processes is N=6. Similar designs involving more than two processes (N>2) under investigation are typical in visual search and some decision making studies.

The main motivation behind extending the SFT approach to testing mental networks of larger sizes was to expand the application in different domains of cognitive tasks. The general extension of the SFT to an arbitrary number of mental processes has been published and detailed in recent work (Yang, et al. 2014).

***N*-Factorial SIC for homogeneous systems: Advances to Higher Factorials**

The present section summarizes the extension of SFT method to large size networks (N>2) that are considered to be the homogeneous systems of mental networks. Homogeneous mental networks are here defined as a set of processes that are organized under a single processing order (serial or parallel) and under a single stopping rule (terminating or exhaustive). For example, Sternberg's serial short-term memory processing model (Sternberg 1967; Ratcliff, 1978) is a homogeneous mental network employing only one type of processing order and one stopping rule for all elements. Ratcliff's (1978) parallel model of memory retrieval is a homogeneous mental network as well [3].

By definition, any mental network made of two processes (N=2) is a homogeneous network as long as there is only one processing order and one stopping rule that could be used. Adding at least one more process to an N=2 network could lead to multiple processing orders and stopping rules.

Statistical tests, the SIC general form

In both MIC and SIC statistics, the N-order difference function over the set of variables of interest is denoted as Δ^n.

[3] In a typical standard short-term memory task (Sternberg, 1969; Ratcliff, 1978) the trials are divided so that there is a total of half target-present and another half of target-absent trials. The target present-condition could employ both self-terminating search rule. That is a subject would stop searching for the target as soon as it is found. In the target-absent condition, all trials must be searched, employing the exhaustive stopping rule to make the correct rejection. Both proposed models of short-term memory search (Sternberg, 1969; Ratcliff, 1978) assumed that a single stopping rule is used in either of the conditions (target-present of target-absent) and that the single processing order was employed. Thus the proposed models are the variants of the class of the homogeneous mental networks.

The N-order factor difference at the level of survivor functions is defined as:

$\Delta S^N(t_{X_{1...N}}; x_{1...N})$

where $t_{X_{1...N}}$ represents a set of response time variables from 1 to N.

This N-order difference function is also denoted as $SIC^N(t) = \Delta S^N(t_{X_{1...N}}; x_{1...N})$ as the SIC function can be generalized to the case for arbitrary N processes. The SIC functions for N=2, 3 and 4 processes are derived below:

Second-order difference:

$SIC^2(t) = \Delta^2 S(t_X, t_Y; x, y) = [S_{LL}(t) - S_{LH}(t)] - [S_{HL}(t) - S_{HH}(t)]$

Third-order difference:

$SIC^3(t) = SIC^2(t; \{L,H\}) = SIC^2(t; L) - SIC^2(t; H) =$
$[S_{LLL}(t) - S_{LLH}(t)] - [S_{LHL}(t) - S_{LHH}(t)] -$
$\{[S_{HLL}(t) - S_{HLH}(t)] - [S_{HHL}(t) - S_{HHH}(t)]\}$

Fourth-order difference:

$SIC^4(t) = SIC^3(t; \{L,H\}) = SIC^3(t; L) - SIC^3(t; H) =$
$[S_{LLLL}(t) - S_{LLLH}(t)] - [S_{LLHL}(t) - S_{LLHH}(t)] -$
$\{[S_{LHLL}(t) - S_{LHLH}(t)] - [S_{LHHL}(t) - S_{LHHH}(t)]\} -$
$\{[S_{HLLL}(t) - S_{HLLH}(t)] - [S_{HLHL}(t) - S_{HLHH}(t)] -$
$\{[S_{HHLL}(t) - S_{HHLH}(t)] - [S_{HHHL}(t) - S_{HHHH}(t)]\}\}$

Figure 1A shows the SIC functions for different mental networks up to the sizes of four processes (N=4). The SIC shapes for the higher order mental networks that are larger than four process can be derived by the induction from the related theorems (Yang, et al. 2014). As it can be seen from Figure 1A, the SIC signatures show a remarkable regularity in changing their shapes as the number of processes under investigation increases. The serial minimum time and the parallel terminating processing networks both show self-repeating patterns as a function of the size of a mental network (N). The serial minimum time SIC signature remains a flat function while the parallel-terminating SIC signature remains a positive unimodal function. These two SIC signatures could be characterized almost as having fractal properties as the output of these functions would recurrently generate the same SIC patterns regardless of the size of the network. To some extent, the self-repeating patterns are also evident in the case of parallel exhaustive system: the SIC function predicts the same shape, albeit the function flips around the x-axis as with each new process (N). Thus, SIC is strictly negative for an even-number of processes and is strictly positive for an odd numbers of processes under investigation.

The serial-exhaustive network shows the distinct coiling behavior around the principle x-axes: for the simplest serial network of the two process, the SIC function is first negative then positive under the condition that the two areas are equal. The serial exhaustive processing for N=2 intersects the x-axes only once, thus introducing diagnostic property of N-1 zero-crossings. By adding more processes, the SIC function flips around the x-axis and coils once more around the x-axis, thus exhibiting another regular shape change "flips and coil" while the total sum of positive and negative areas is always zero.

The results of the recent extension of SFT to multiple processes (Yang, et al. 2014) provides new opportunities for extend exploration of fundamental properties of larger mental networks, and it provides means for validation/falsification operations in various cognitive tasks.

Limitations

The diagnostic signatures SIC derived for homogeneous mental networks (Yang, et al. 2014) are subjected to the two general limitations that can lead to diminishing the SFT diagnostic power: (1) Non-unique SIC patterns: Upon visual inspection of Figure 1A it could be observed that there are some cases of different mental architectures predicting the same shape of the SIC function. Some mental networks are made of different fundamental mental properties which have an identical SIC signature for the same *N*'s. For example, Figure 1A shows that both the parallel minimum time and the parallel exhaustive network make the same SIC signature prediction when the number of processes is odd (N=3, 5, 7 …) [4].

(2)The diagnostic SIC signatures derived for homogeneous mental networks cannot be generalized to all types of mental networks. One such case is the class of non-homogeneous networks of larger sizes (N>2) which combine different processing properties within one mental network. For example, a case of a non-homogeneous network has some mental operations conducted in serial and some conducted in parallel[5].

In the current chapter, I provide new evidence and insights that could be used to address the concerns raised by the apparent limitations of the current SFT methods on the exploration of homogeneous systems consisting of more than two processes (N>2). These efforts could be seen as another extension of the SFT methods and as a suggestion

[4] Self-repeating SIC patterns across different sizes N: For some mental networks, the SIC signatures appears of identical shape across different mental network sizes. For example, the unimodal positive SIC function which is used to indicate the presence of the parallel minimum time processing network could not be used to differentiate whether the system used 2, 3, 4 or N number of processes. The self-repeating patterns could be characteristic of this function of having fractal properties. The solution for the self-repeating SIC patterns is provided by the experimental method. The size of the network is manipulated by an experimenter and is specified in the experimental method.

[5] Another example would be the case of probability mixtures of homogeneous mental networks such that a subject would switch from one type of homogeneous mental network to another type on each new trial. For example, switching from pure serial to pure parallel in repeated experimental trials. Although the class of probability mixtures of the homogeneous mental network is an important case, it will not be covered in the current chapter.

about the new approaches which test even more complex mental networks. In order to unlock potentially more powerful diagnostic features of SFT approach, I will integrate some previous approaches with the novel ideas. The ideas will be reported in the form of proposals and their proofs in the following section. Within some of the proposals, I will provide the analysis of illustrative cases.

Simple-Factorial SIC functions for homogeneous systems

The existence of non-unique SIC signatures reduces the diagnostic power of SFT of homogeneous systems which is made of arbitrary number of processes (N>2). Note that each SIC function in Figure 1A is obtained by the analysis of the full factorial SFT design. That is, if the number of processes analyzed in the underlying mental network was *N*, then *N* variables were engaged in the full-factorial SFT design. I argue here that the application of the lower-level interaction tests unlocks the additional diagnostic power of SFT tools. The argument is based on the findings that the simple interaction tests would provide more evidence to distinguishing between different types of mental networks.

Claim: Applying a simple interaction factorial analysis on a full-factorial data set improves the diagnosticity of the SFT when applied on homogeneous mental networks of sizes larger than two under the conditions specified.

Evidence: One can derive the new sets of SIC signatures for the lower-level factorial designs within a full factorial design. This is achieved by marginalization of the effects of some factors in a full-factorial design and by inspection of the SIC signatures obtained. For example, a four-process network, that is investigated by the four factors full-factorial design, can also be investigated by the four three-way factorial designs and six two-way factorial designs by the means of marginalization of single factors.

The procedure of deriving low-level interactions is equal to the one conducted in factorial ANOVA. Marginalization[6] is equal to averaging out the effect of unwanted factors. Given that a factor in SFT design represents a single process of interest, it could be in either a low or high state (selectively influenced by the experimental manipulation). Its marginalization would lead to excluding this factor from the design by averaging its effect across other factors.
In practice, marginalization of the factors leads to having to combine the result conditions of the high and low stretching into a single condition. Assume a network is made of three processes (X,Y and Z) then the third order difference function is shown in a simpler form:

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

Each three subscript letters represent a different variable (X, Y, Z). Each variable represents a single process that could be in either a low (L) or high (H) state based on stretching manipulation. In the experimental study, each term represents an empirical survivor function of one experimental condition consisting of a sample of repeated RT trials.

To derive the low level-interaction contrast functions of the second-order difference, one has to marginalize the effect of one variable in the equation. To find the second-order difference of the SIC function of the two processes between Y and Z, one has to marginalize the effect of X**.** To marginalize the variable X, which is the first variable in the subscript, one has to aggregate the low and high conditions for the first variable which leads to the following second-order difference between the two variables, Y and Z:

$$SIC^{2\,(\cdot,\,Y,\,Z)}(t) = [S_{LLL+HLL}(t) - S_{LLH+HLH}(t)] - [S_{LHL+HHL}(t)- S_{LHH+HHH}(t)]$$
$$=[S_{.LL}(t) - S_{.LH}(t)] - [S_{.HL}(t) - S_{.HH}(t)] \quad (5)$$

The dot indicates a marginalized factor. The plus sign means the operation of union of data sets from the two different conditions (LLL+HLL). A similar procedure is used to find the other two second-order difference SIC functions ($SIC^{2\,(X,\,\cdot\,,Z)}$ and $SIC^{2\,(X,\,Y,\,\cdot)}$. We will refer to these as to the <u>simple interaction SIC tests</u>.

This form (Eq 5) of the simple interaction SIC test above is equivalent to the form of the full-factorial interaction test. The second row in Eq 5 is the result of the first variable marginalization. This outcome is equivalent to the second-order difference $SIC^2 = [S_{LL}(t) - S_{LH}(t)] - [S_{HL}(t) - S_{HH}(t)]$. One can show that derivation of the simple interaction SIC test of any factorial design will lead to the equation forms that are equivalent to the full-factorial SIC forms lessened by the number of marginalized factors.

By using simple interaction SIC tests, and marginalization of variables, it is possible to derive all possible lower-order differences for any full-factorial design. The relevant SIC predictions for each distinct mental network, of a size up to four processes, are presented in Figure 1B.

The analysis of **simple interactions** shown in Figure 1B directly addresses the issue raised by the non-unique SIC pattern limitation when the SFT full-factorial designs were used. For example, in the case of shared SIC signature between the minimum-time parallel model and the parallel exhaustive model for the odd number of process (N=odd) under an investigation, the simple interaction factorial analysis now provides the set of simple SIC interaction contrast functions for each full-factorial design (Figure 1B). For example, take the case of N=3 SIC signatures, presented in Figure 2. Although the third-order

[6] Another approach would be to condition rather than to marginalize but it is not addressed in this work.

difference SIC functions are shared (thus non-unique) between parallel minimum time and parallel exhaustive models, the derived simple interactions (two-way) show the distinct SIC patterns. In the case of parallel minimum time model, the 2-way simple SICs are all positive while for the exhaustive parallel model the correspond 2-way simple SICs are all negative.

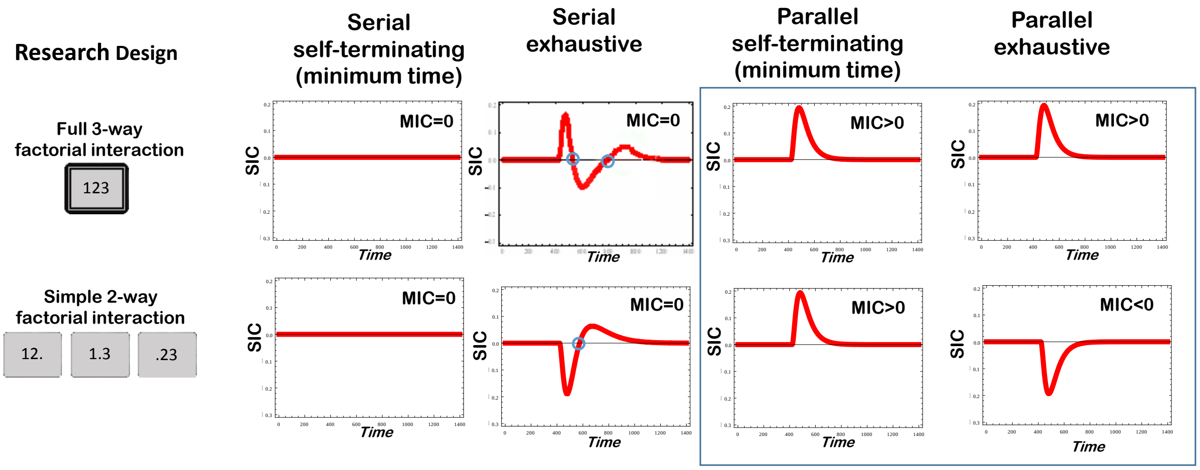

Figure 2: The simple interaction SIC test for the 3-factorial research design. The research starting with the full factorial are displayed in the first column, crossed with the different types of mental networks. Although the full-factorial SIC cannot distinguish between parallel minimum time and parallel exhaustive models (the first row, SICs of the models are in a box), the simple two-way factorial SIC (the second row, boxed) can distinguish between the two.

Limitations

The SFT methodologies described so far are confined to exploration of homogeneous mental networks. The examples of the homogeneous networks are pure serial or pure parallel processing systems in which one stopping rule (either OR or AND) is applied on all network processes. Unfortunately the SFT signature predictions would no longer hold valid if a non-homogenous mental networks (such as a serial-parallel network) is analyzed.

Although the research literature proposes many homogenous mental network (e.g. Sternberg, 1967; Ratcliff, 1978) many researchers have provided the evidence which shows that it is not realistic to assume that the fundamental processing properties are always homogeneously distributed across larger mental networks. In fact, it is plausible to assume that larger mental networks could combine several fundamental properties within a single mental network.

***N*-Factorial SIC for non-homogeneous networks**

One formal way to describe a cognitive system that can combine different fundamental properties is a directed acyclic network (Schweickert, et al., 2000). For example, the networks depicted in Fig 4, the first row, combines both serial and parallel processing. These mental networks utilize either OR or AND stopping rule. This means that the network will wait for the completion of both items before proceeding to the next stage (AND=exhaustive processing), or will terminate on completion of a single process (OR=processing termination is possible).

Take for example a model of short-term memory. The original Sternberg’s model was later challenged by the results that showed a strong serial position recency effect in memory scanning. That is, recently stored items were analyzed faster than the older items in the set. The recency effect suggests that recent items may be stored in a different way than the older items. To account for the recency effect, several alternative models were published which proposed the idea that short-term memory processing consists of two subnetworks, serial and parallel. It was proposed that short term memory consists of two distinct temporal stores, so the items stored can have different accessibility rate (Clifton & Birenbaum, 1970; Waugh & Norman, 1965; Posner & Taylor, 1969; Burrows & Okada, 1971; Forin & Cunningham, 1973; see also Oberaurer, 2002; Oberauer & Bialkova, 2009)

Work on the identification of the acyclic serial-parallel networks has been conducted by Richard Schweickert, Ehtibar Dzhafarov, and a group of collaborators (Schweickert, Fisher & Goldstein, 2010; Schweickert, et al., 2000; Dzhafarov, Schweickert, & Sung, 2004; for an overview see Schweickert, Fisher & Sung, 2012). I will summarize the main findings concerning the SIC predictions which regard the serial/parallel networks.

Statistical tests and subnetwork decomposability

When an underlying processing system is nonhomogeneous, one can expect that the N-factorial SIC test will become less diagnostic with the increase of the number of processes and the level of heterogeneity. For example, imagine a large size network of up to 4 processes in which some of them form serial and some form parallel subnetworks, each of which could utilize different stopping rules. Let us assume that the described 3-process network is a directed acyclic network not known to researchers. The researchers' task is to reveal the network in terms of the fundamental mental properties (processing order, stopping rule, dependency).

Theoretically, the 3-process network, which could be either serial or parallel, can be organized in many ways. If the mental network is homogeneous then there is only a single solution about how to organize them, by using one processing order and one stopping rule. If the 3-process mental network is non-homogeneous then that number could be rather high: the number of different ways to organize them is close to one-hundred combinations.[7]

Although the SFT analysis could be employed at the full-factorial level analysis, it is immediately striking that such a high number of possible mental network organizations would generate too many three-factorial SIC predictions. In the best case, some of these 3-process combinations would generate unique SIC patterns that could be used to clearly distinguish them from other possible combinations. However, it is more likely that many such combinations would produce identical SIC shape predictions, thus with the potential to diminish applicability in differentiation of subnetworks.

In the above section on the sample interactive factorial tests, we learned that more diagnostic power is achieved by analysis of the lower level factorial interactions. The SFT analysis of the nonhomogeneous system should proceed from the full factorial to a simpler factorial analysis. In fact, the strategy of factoring out (dropping) some processes from the full factorial design proves to be effective when applied in SFT factorial analysis. As in the above homogeneous network analysis case (Figure 2), the marginalization of the effects of some processes allows us to reach the simplest subnetworks, consisting of only two processes (N=2), which are the homogeneous networks by definition. The converging idea here is that a simple-factorial analysis should improve diagnosis of possible different subnetworks within a more complex nonhomogeneous one, thus helping to improve the diagnostic power of the SFT analysis.

An immediate concern here is that even though we can isolate and analyze small-size subnetworks within a more complex one, what SIC predictions should we expect to see? For example, a two-process serial subnetwork may be embedded within a system of parallel process inside the entire network. The question here is whether the identification of such a small serial subnetwork would be affected because of the connection to other heterogeneous parts of network? This concern is reasonable, because the observed data about the small size network is based on the response of the entire mental network.

A large body of such a work has already been conducted (Schweickert, Giorgini & Dzhafarov, 2000; Dzhafarov, Schweickert, & Sung, 2004) and it is possible to summarize. The work is limited to the so called serial parallel-networks in which a single dominant stopping rule is used, but the processes can be organized in either serial, parallel or combined subnetworks (e.g. Schweickert, et al., 2000). The general theorems apply to many cases in a distribution free manner.

Findings

Under the same conditions specified above the summary of the SIC function expectations is presented below and in Figure 3. In the case of detection of a parallel subnetwork (N=2) connected to a serial process(es) within a larger network, with either AND or OR gates, the SIC functions predictions are identical to those of an isolated parallel two-process network (Theorem 4, case 1, from Schweickert, et al., 2000). That is, the diagnostic SIC shape of the parallel subnetwork does not change its shape with the presence of another serial process within the same network (in Figure 3). Thus, in the case of parallel AND gate the SIC function is negative for all times. In the case of parallel OR gate, the SIC function is positive for all times.

In contrast, predictions for the serial subnetworks are less specific than those for the parallel subnetworks. In the case of **serial subnetwork** (N=2) connected to a parallel process(es) within a larger networks, the SIC functions are not identical to those of isolated two-process serial networks. In the AND serial subnetwork (N=2), the area under the SIC function is negative for a short period of time (Theorem 5, Schweickert, et al., 2000; Theorem 6.2 Dzhafarov, et al., 2004) and then changes sign at least once. Overall, the area under the SIC curve is equal or larger than zero (Theorem 6.1 from Dzhafarov, et al., 2004) (Figure 3). In the OR serial subnetwork (N=2), the area under the SIC curve is equal or lower than zero (Theorem 5, Schweickert, et al., 2000; Theorem 6.1 from Dzhafarov, et al., 2004) (Figure 3)

Limitations

The theorems and proofs (Schweickert, et al., 2000; Dzhafarov, et al., 2004) provide the general diagnostic shape of SIC function for a small size (N=2) mental networks embedded in a larger serial-parallel network. The prediction results are of limited diagnostic power for detecting unknown subnetworks. The reason is that in some cases different types of subnetworks would predict the same SIC under the general specifications and the listed conditions. The formal proofs operate at the level of non-strict inequalities. For example, such approach cannot be used to distinguish between the serial AND subnetwork and the parallel OR subnetwork and between the serial OR subnetwork and the parallel AND subnetwork. These two pairs cannot be disentangled given the SIC results in an unknown underlying mental network.

[7] To get the correct number of possible combinations of organization of 3 processes, one has to take into account that each process can be either serial or parallel (8 combinations for 3 processes) and that each process could be assigned either the OR or AND stopping rule (another 8 combinations), thus making a total of 8 x 8=64 combinations.

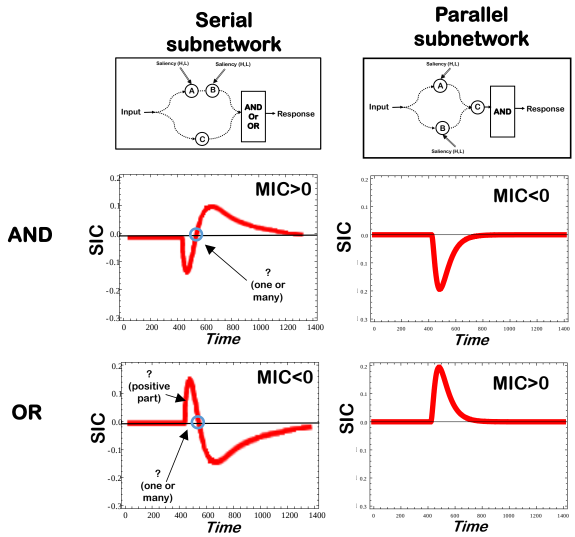

Figure 3: SIC functions predictions for the serial and parallel subnetworks, cross with the minimum time stopping rule (OR) and exhaustive stopping rule (AND). The factorially manipulated (indicated by the arrows in the first row) processes, are embedded in the serial-parallel network. The SIC shape expectations are not exact because they are the result of not so strict inequalities (Schweickert, et al., 2000; Dzhafarov, et al., 2004). The unknown properties are indicated by the question signs

Putting it all together: homogeneous and non-homogeneous subnetworks *N*=2

It is still possible to recuperate more of the diagnostic power (Theorem 5, Schweickert, et al., 2000; Dzhafarov, et al., 2004) from an SFT analysis. The SFT signatures are shared between different serial and parallel subnetworks which utilize different stopping rules (AND and OR). In many cases in research studies, it is possible to fix the stopping rule methodologically. Such examples are target-present or target-absent responses which are designed by a researcher. Providing high accuracy in subjects' responses, a researcher is able to analyze target-present responses as likely candidates of using the OR rule and target-absent response as using the AND rule. In other words, a researcher can separate experimental conditions in which the AND rule or OR rule are used. If the stopping rules are fixed by researcher, and analyzed separately by target-present and –absent conditions, then the SIC signatures can be used to distinguish between the serial and parallel processing subnetworks, as depicted in Figure 3.

The question is why do parallel subnetworks (*N*=2) embedded in a serial network predict the same SIC results as a homogeneous parallel network of the same size (N=2) while, at the same time, the serial system subnetwork (*N*=2) embedded in a parallel processing network doesn't predict the same SIC results as a homogeneous serial network of the same size?

Consider the case depicted in Figure 3, left column, which show that the two serial processes, A and B, are combined in parallel with the third process C. If A and B are selectively influenced by the corresponding factorial stretching manipulations (to produce HH, HL, LH and LL) then what should be the predicted SIC function for the two serial processes?

The proof (Schweickert, et al., 2000; Dzhafarov, et al., 2004) shows that, depending on the stopping rule, the SIC function for the two serial processes (t_A+t_B) should be zero or largely positive if the AND gate is used as the stopping rule, that is, the system waits for the slower of the two components, that is Max(t_A + t_B, t_C). The SIC function should span the areas below and above x-axis that is ether zero or is negative (it can have positive values too, but the total sum of the spanned area is negative). If the OR gate is used, that is, the system stops on the first completed component Min(t_A+ t_B, t_C), the approximate SIC shape is not really known: the number of x-crossings of SIC function is not known. What is known is that the total area spanned by the SIC is either positive (AND) or negative (OR). In other words, the corresponding MIC values should be positive (AND) or negative (OR).

An alternative perspective into the result of SIC function can be used by analyzing the MIC values of the double factorial difference on the situation depicted in Figure 3, the first subnetwork (AND) on the left.

For example, for the subnetwork $\mathrm{Max}(t_A + t_B, t_C)$, it is possible to state the following:

$$E[\mathrm{Max}(t_A + t_B, t_C); \mathrm{a,b,c}] = p \cdot E[t_A + t_B; \mathrm{a,b}] + (1-p) \cdot E[t_C; \mathrm{c}]$$

That is, the total time to complete processing in the serial-parallel network depicted in Figure 3, left panel, is equal to the probability mixture of the two events: One, when the slowest component is $t_A + t_B$ with probability *p*, and two, when the slowest component coming from the parallel process t_C with probability (1-*p*).

Another observation is that the random variable in the underlying process C has a fixed rate (not stretched) while the rates for A+B depend on the factorial stretching effects imposed on processes A and B (a,b={low and high})

In the parallel system, probability *p* is defined as the probability that process A+B has completed before the second parallel process C. Here, it is shown as the integral over the completion density function of t_A+t_B (that is, t_A+t_B is completed first, thus the subscript 1), and the survivor function specifying that the process C has not been completed yet (has not completed first, thus the subscript 1).

$$p < AB, \mathrm{C} >= \int_0^\infty f_{AB_1}(t) \cdot S_{C_1}(t) dt$$

Since we are focused on the sign of the MIC value (that reveals the area under the SIC function) of the two serial processes A and B,

$$\Delta^2 E[\mathrm{t} = \mathrm{Max}(t_A + t_B, t_C); \mathrm{a,b,c}] =$$
$$\Delta^2\{p \cdot E[t_A + t_B; \mathrm{a} = \{\mathrm{low,high}\}, \mathrm{b} = \{\mathrm{low,high}\}] + (1-p) \cdot E[t_C; \mathrm{c}]\} \Rightarrow$$

$$\begin{aligned}
&\mathrm{MIC}(\mathrm{t} = \mathrm{Max}(t_A + t_B, t_C)) = \\
&p_{LL} \cdot E[t_A + t_B; LL] + (1-p_{LL}) \cdot E[t_C; \mathrm{c}] - \\
&p_{LH} \cdot E[t_A + t_B; LH] + (1-p_{LH}) \cdot E[t_C; \mathrm{c}] - \\
&p_{HL} \cdot E[t_A + t_B; HL] + (1-p_{HL}) \cdot E[t_C; \mathrm{c}] + \\
&p_{HH} \cdot E[t_A + t_B; HH] + (1-p_{HH}) \cdot E[t_C; \mathrm{c}]
\end{aligned}$$

One can immediately infer that the rate of process C would directly affect the MIC sign through the value of the parameter *p*. One can expect that the probability of the processing order *p* will change depending on the relative speeds of the parallel processes A+B and C. For example, in the AND system, if the process C is very fast (*p* is approaching to 1), then the system will stop more frequently on the serial A+B completion times; if the process C is very slow (*p* is approaching to 0) then the system is more likely to stop on completion time of the process C. It is interesting observation that in this case the MIC value is a probability mixture of different factorial conditions (*HH, HL, LH and LL*) which are unequally weighted with the probability *p*. That is, the p value is different for each factorial condition (p_{LL}, p_{LH}, p_{HL}, p_{HH}) depending on the relative rate of processing completion of both the A+B and C components. We know from the proof of Schweickert, et al. (2000) that the SIC function for such a mental network is mainly positive function thus the MIC>0.

In order to observe the shape of the SIC function of such a mental network, I provided the converging evidence by simulation. The results are based on the extensive simulations of the random walk process with absorbing boundaries which are used to characterize rate of completion of each process in the subnetwork. The simulations were conducted across different values of parameters. Here, the converging result is that, in the AND serial subnetwork, the SIC is S-shaped function that have only one x-crossing with a larger positive area then negative area (MIC>0) (see Figure 4, the first row). The simulation results add to the converging evidence to the simulation using the class of exponential distribution (Schweickert, et al., 2000; Dzhafarov, et al., 2004). The simulation results in Figure 4 show that depending on the relative speed of the parallel processes, the SIC shapes undergo the expected transformation: as the process C becomes faster, then the revealed SIC and MIC shape indicate pure serial N=2 mental network. This is because as C becomes very fast then the network finished mostly on the completion of the serial processes A+B.

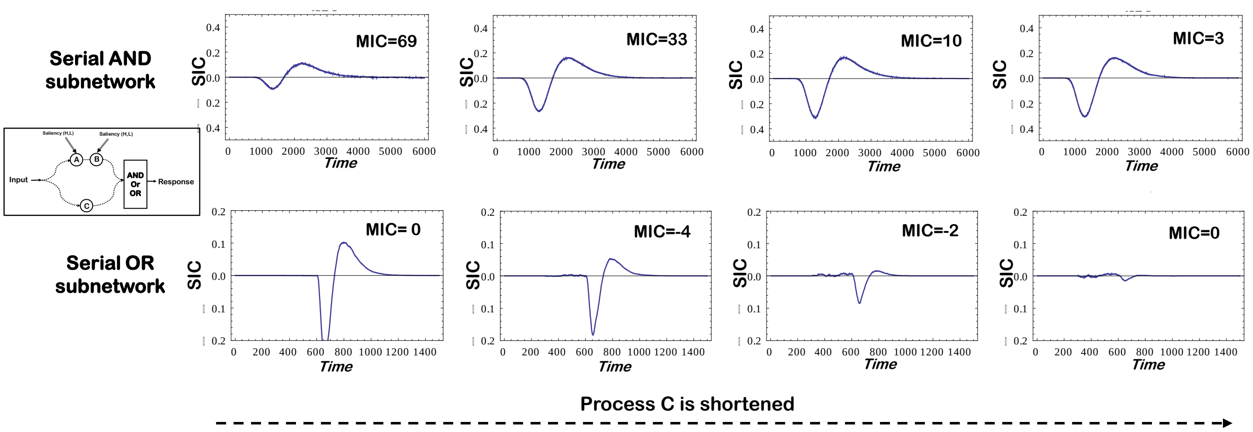

Figure 4: The simulation results of the serial-parallel network depicted in the first column for two different stopping rules (AND and OR). The subnetwork of interest is made of two serial processes and one in parallel with the first two. Across columns, the duration of the third parallel process is manipulated, such that the process is shortened, while the two processes in serial are of fixed parameter value time duration. In the simulation model each processing time completion was determined by the simple random walk process with two bounds and probability *p* of stepping to one of the bounds. 10000 trials were conducted per factorial condition.

The simulation results for Serial OR subnetworks (Figure 3, left panel the second row), showed one crossing S-shaped SIC function, similar to those of Serial AND subnetwork. As shown in a series of simulations based on the duration of the parallel process, the faster it becomes, it is more likely that the subnetwork will finish on a parallel process first and would not wait for the completion of the two serial processes. The result is that the SIC shape becomes more squished, showing some small negative MICs, until it completely dissolves into a straight line.

In contrast, the serial-parallel AND network depicted in Figure 3 (second column) shows the parallel subsystem embedded in the serial network. The predicted MIC value and SIC function are equivalent to the pure (homogeneous) parallel AND model for the N=2 number of processes. (The proofs are straightforward and is presented in Schweickert, et al., 2000, p. 502, Case2; and also in Fific, 2006, Appendix). Also when the OR gate is used in the same network both the MIC and SIC make predictions as the pure (homogeneous) parallel OR model for the N=2 number of processes.

Discussion

There has been great deal of progress over the last fifty years in both the theoretical development of various cognitive models and in developing methods of validation of such models. The building blocks of cognitive models are defined as the fundamental properties of mental processes. These are: processing order, stopping rule, process dependency, and processing capacity.

The current dominant approaches to exploring underlying cognitive models rely mostly on formal parametric descriptions of the cognitive models. So, the cognitive models are seen as a family of a finite number of distributions that are described by a set of parameters. Consequently, cognitive models have been tested by examining the parameter values, observing the parameter values that are the most likely to generate data (least-square and maximum likelihood estimations), by applying the model selection procedures (AIC, BIC), or by using more advance techniques to account for prior parameter values such as in the Bayesian model analysis (for a short review see Liu, & Smith, 2009).

Such approaches provide an invaluable set of tools for model exploration. However, these approaches depend on the parametric assumptions of the models. In many cases, the model exploration is carried over by the very time consuming and computer intensive methods. In order to find the best fitting model's parameters, one has to employ optimization search for parameters, which are, even nowadays, limited by computational power. Even the most advance methods, such as Bayesian inference and model selection (Lee & Wagenmakers, 2013; Raftery, Gelman, Rubin, & Hauser, 1995) require a simulation method known as Markov chain Monte Carlo (MCMC) which is not immune from days of sampling of simulated data from the unknown parameter values.

The groundbreaking work of James Townsend and other essential contributors led to the creation of the Systems Factorial Technology (SFT) - a suite of methodological tools for directly investigating the fundamental properties of cognitive operations. The SFT approach, created by James Townsend, rests on rigorously tested mathematical tools for discriminating between serial and parallel processing, exhaustive and self-terminating stopping rules, and stochastic independence and dependence, as well as for discerning the capacity of an investigated system, all in a parameter-free manner.

SFT is an alternative to exploring a parameter space for the process of interest. The SFT requires factorial research design of N-number of binary valued factors that selectively stretch the processes of interest, and requires the response time measure. The response time results are either aggregated (mean) or used to estimate the survivor functions, across the respective factorial conditions. The N-order difference function is applied on the results and the two statistics are obtained MIC and SIC. The MIC and SIC are examined to make inference about the fundamental properties of processes. The stronger statistic here – the SIC function – reveals the different signatures for different fundamental properties.

The present study is focused on further refining recent advances in SFT methodology (Yang, Fific & Townsend, 2014) and on the development of new tools for use with larger mental networks. The motivation for this study was to address and remove the two limitations of the current SFT methods when applied on larger mental networks. The first concern was revealed when SFT was applied on the increasing number of processes under investigation (Yang, Fific & Townsend 2014): the predicted SIC signatures were shared between different cognitive models. This concern was by addressed by inclusion of simple interaction SIC tests conducted on N=2 subnetworks, the simple-interaction SIC test were carried over by ignoring some variables in a higher order factorial design (N>2) and dropping them from the factorial design. The second concern was that the main results so far have been confined to the class of the homogeneous mental networks and neglect the possibility that underlying mental networks is non-homogeneous (such as serial/parallel networks)

To address this concern the present study integrated the results of the simple interaction SIC analysis for the higher order factorial design (N>2) with the factorial tools developed to explore non-homogeneous mental networks which may consist of both serial and parallel processes (so-called serial/parallel networks Schweickert, et al., 2000; Dzhafarov, Schweickert, & Sung, 2004). The results of the integration are summarized so that the SIC signature expectations for various N=2 subnetworks were generated. The current study also provided the converging evidence from the simulations regarding the detailed SIC expectations.

As a strong alternative to parameter-dependent model-testing approaches, the SFT is a powerful tool to analyze processes underlying any cognitive activities. This study calls for further exploration of more complicated serial-parallel mental networks and further studies that should extend to current ongoing revolution in the analysis of neural networks (e.g., Lisman, & Idiart, 1995; Verwey, Shea, & Wright, 2015; Agliari, Barra, Galluzzi, Guerra, Tantari, & Tavani, 2015; Rushworth, Kolling, Sallet, & Mars, 2012; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Woodman, & Luck, 2003; Pooresmaeili, Bach, & Dolan, 2014; Ward, 2003; Raghavachari, Kahana, Rizzuto, Caplan, Krischen, Bourgeois, Madsen & Lisman, 2001; For the review of the current cognitive methods borrowed by the neural approaches see Caplan, 2009). Future studies on how to use SFT could provide an important window in the organization of mental process and model validation procedures which would not be easily paralleled by the other approaches.

References

Agliari, E., Barra, A., Galluzzi, A., Guerra, F., Tantari, D., & Tavani, F. (2015). Hierarchical neural networks perform both serial and parallel processing. *Neural Networks, 66*, 22-35.

Anderson, M. J., & Whitcomb, P. J. (2000). DOE simplified: Practical tools for effective experimentation. Portland, Or: Productivity.

Burrows, D. & Okada, R. (1971). Serial position effects in high-speed memory search. *Perception & Psychophysics*, *10* (4B), 305-308.

Caplan, D. (2009). Experimental design and interpretation of functional neuroimaging studies of cognitive processes. *Human Brain Mapping, 30*(1), 59-77

Clifton, C. & Birenbaum, S. (1970). Effect of serial position and delay of probe in a memory scan task. *Journal of Experimental Psychology*, Vol. 86, No. 1, 69-76.

Dzhafarov, E. N., Schweickert, R., & Sung, K. (2004). Mental architectures with selectively influenced but stochastically interdependent components. *Journal of Mathematical Psychology*, *48*(1), 51-64.

Donders, F.C. (1868). La vitesse des actes psychiques. *Archives Néerlandaises des Sciences Exactes et Naturelles, 3,* 296-317.

Eidels, A., Townsend, J. T., & Pomerantz, J. R. (2008). Where similarity beats redundancy: The importance of context, higher order similarity, and response assignment. *Journal of Experimental Psychology: Human Perception and Performance, 34(6),* 1441-1463.

Egeth, H., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 551–560.

Forrin, B. & Cunningham, K. (1973). Recognition time and serial position of probed item in short-term memory. *Journal of Experimental* Psychology, Vol. 99, No. 2, 272-279.

Fific, M. (2006). *Emerging holistic properties at face value: Assessing characteristics of face perception.* Unpublished doctoral dissertation, Indiana University, Bloomington.

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.

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(2),* 356-375.

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., & Townsend, J. T. (2010). Information-processing alternatives to holistic perception: Identifying the mechanisms of secondary-level holism within a categorization paradigm. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 1290-1313.

Fific M (2014) Double jeopardy in inferring cognitive processes. Front. Psychol. 5:1130. doi: 10.3389/fpsyg.2014.01130

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science, 7*(2), 99-108.

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

Houpt, J.W., Blaha, L.M., McIntire, J.P., Havig, P.R. and Townsend, J.T. (2014). Systems Factorial Technology with R. Behavior Research Methods, 46, 307-330.

Johnson, S. A., Blaha, L. M., Houpt, J. W., & Townsend, J. T. (2010). Systems factorial technology provides new insights on global–local information processing in autism spectrum disorders. *Journal of Mathematical Psychology, 54*(1), 53-72.

Lee, M. D., & Wagenmakers, E. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press, New York, NY.

Lewis-Peacock, J., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience, 24*(1), 61-79.

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

Little, D. R., Nosofsky, R. M., Donkin, C. & Denton, S. E. (2013). Logical-rules and the classification of integral dimensioned stimuli. *Journal of Experimental Psychology: Learning, Memory & Cognition, 39,* 801-820.

Liu, C. C., & Smith, P. L. (2009). Comparing time–accuracy curves: Beyond goodness-of-fit measures. *Psychonomic Bulletin & Review, 16*(1), 190-203.

Lisman, J. E., & Idiart, M. A. P. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science, 267*(5203), 1512-1515.

Maxwell, S. E., & Delaney, H. D. (1999). Designing Experiments and Analyzing Data: A Model Comparison Perspective. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Pooresmaeili, A., Bach, D. R., & Dolan, R. J. (2014). The effect of visual salience on memory-based choices. *Journal of Neurophysiology, 111*(3), 481-487.

Posner, M. I., & Taylor, R. L. (1969). Subtractive method applied to separation of visual and name components of multiletter arrays. *Acta Psychologica, Amsterdam, 30*, 104-114.

Raftery, A. E., Gelman, A., Rubin, D. B., & Hauser, R. M. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111-184.

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Krischen, M. P., Bourgeois, Madsen, J. R., Lisman, J. E. (2001). Gating of human theta oscillations by a working memory task. *The Journal of Neuroscience, 21*(9), 3175-3183.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108. doi:http://dx.doi.org/10.1037/0033-295X.85.2.59

Rushworth, M. F. S., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: One or many serial or parallel systems? *Current Opinion in Neurobiology, 22*(6), 946-955.

Schweickert, R., Fortin, C., & Sung, K. (2007). Concurrent visual search and time reproduction with cross-talk. *Journal of Mathematical Psychology*, *51*(2), 99-121.

Schweickert, R. (1999). Response time distributions: Some simple effects of factors selectively influencing mental processes. *Psychonomic bulletin & review*, *6*(2), 269-288.

Schweickert, R., & Dzhafarov, E. (2000). Selective influence and response time cumulative distribution functions in serial-parallel task networks. *Journal of Mathematical Psychology*, *44*(4), 504-535.

Schweickert**,** R.**,** Fisher, D. L., & Sung, K. (2012). Discovering cognitive architecture by selectively influencing mental processes. Singapore: World Scientific.

Schweickert, R., Fisher, D. L., & Goldstein, W. M. (2010). Additive factors and stages of mental processes in task networks.*Journal of Mathematical Psychology, 54*(5), 405-414.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652- 654.

Sternberg, S. (l969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist, 4,* 421-457.

Sung, K. (2008). Serial and parallel attentive visual searches: Evidence from cumulative distribution functions of response times. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(6), 1372.

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

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

Townsend, J. T. (1971). A note on the identifiability of parallel and serial processes. *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., & Ashby, F.G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, UK: Cambridge University Press.

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

Verwey, W. B., Shea, C. H., & Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychonomic Bulletin & Review*, *22*(1), 54-77.

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences, 7*(12), 553-559.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, *72,*89-104.

Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 755–779.

Wenger, M. J., & Townsend, J. T. (2001). Faces as Gestalt stimuli: Process characteristics. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges. Scientific psychology series* (pp. 229-284). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Woodman, G. F., & Luck, S. J. (2003). Serial deployment of attention during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 121-138.

Yang, C. (2011). Relative saliency in change signals affects perceptual comparison and decision processes in change detection. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1708-1728. doi:http://dx.doi.org/10.1037/a0024257

Yang, H., Fific, M., & Townsend, J. T. (2014). Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes. *Journal of Mathematical Psychology, 58,* 21-32. doi:http://dx.doi.org/10.1016/j.jmp.2013.12.001

Yang, C., Chang, T., & Wu, C. (2013). Relative change probability affects the decision process of detecting multiple feature changes. *Journal of Experimental Psychology: Human Perception and Performance*, 39(5), 1365-138