Summer August 31, 2017

A hierarchical Bayesian approach to distinguishing serial and parallel processing

Joseph W Houpt
Mario Fifić

A Hierarchical Bayesian Approach to Distinguishing Serial and Parallel Processing

Joseph W. Houpt[a], Mario Fifić[b]

[a]*Department of Psychology, Wright State University, Dayton, Ohio*
[b]*Psychology Department, Grand Valley State University, Allendale, Michigan*

Abstract

Research in cognitive psychology often focuses on how people deal with multiple sources of information. One important aspect of this research is whether people use the information in parallel (at the same time) or in series (one at a time). Various approaches to distinguishing parallel and serial processing have been proposed, but many do not satisfactorily address the mimicking dilemma between serial and parallel classes of models. The mean interaction contrast (MIC) is one measure is designed to improve discriminability of serial-parallel model properties. The MIC has been applied in limited settings because the measure required a large number of trials and lacked a mechanism for group level inferences. We address these shortcomings by using hierarchical Bayesian analyses. The combination of the MIC with hierarchical Bayesian modeling gives a powerful method for distinguishing serial and parallel processing at both individual and group levels, even with a limited number of participants and trials.

Keywords: response times, Bayesian modeling, systems factorial technology, serial parallel

1. Introduction

Situations in which people need to combine several sources of information are ubiquitous. Often, people must switch among cognitive strategies for dealing with these multitudinous sources depending on the situation. Take for example a fighter pilot in combat whose life depends on the successful, simultaneous utilization of several sources of information, i.e., a parallel processing cognitive strategy. In contrast, the same pilot may be required to utilize rather different type of cognitive processing strategy when following up lengthy preparatory flying technical procedures. For example, to turn on plane's engines an operator must usually conduct several operations in a strictly non-overlapping sequence.

URL: joseph.houpt@wright.edu (Joseph W. Houpt)

The failure to stick to the strict sequence of operations may have a fatal consequence. In everyday life, deployment of different cognitive strategies may not be associated with fatal outcomes, but may nonetheless have important consequences.

Given the prevalence of tasks that require multiple sources of information to be attend to, it is no surprise that the properties of the cognitive processes underlying the combined use of those sources of information is a major topic of investigation in modern cognitive science. Cognitive scientists have operationalized the four fundamental cognitive operations for dealing with multiple sources. The first is the temporal organization of the information processing. Processing may be serial, i.e., item-by-item analysis, or parallel, i.e., all-items-at-once. The second is stopping rule, which refers to whether a cognitive system can terminate processing after completion of only a few processes, henceforth referred to as self-terminating, or a system has to complete all processes, henceforth referred to as exhaustive. The third is process interdependency: the extent to which processes of interest depend on each other. The fourth property is processing capacity, which refers to how much processing resources are available for cognitive operations.

Without carefully crafted empirical designs and inferential tools, even processing characteristics as distinct as serial and parallel processing can be perfectly indistinguishable. For example, the standard serial and the limited capacity parallel models cannot be distinguished from each other using the conventional performance measures such as mean response time or accuracy (e.g., Townsend, 1971, 1972; Townsend & Ashby, 1983, Chapter 14).

One framework that has resulted in success at assessing the fundamental properties of cognitive operations is systems factorial technology (SFT; Townsend & Nozawa, 1995; Schweickert et al., 2000; Dzhafarov et al., 2004; Houpt et al., 2014). The SFT approach rests on rigorously tested mathematical tools for discriminating serial from parallel processing exhaustive from self-terminating processing, process (in)dependence and the capacity of the system under investigation. In the current project we focus on discriminating between the parallel or serial processing of two sources of information, however SFT has been generalized to diagnosing underlying system of any number of processes (Yang et al., 2014b; Zhang & Dzhafarov, 2015; Fifić, 2016).

In our current paper, we present methods for inference based on a particular measure from SFT, the mean interaction contrast (MIC). In the case of the mental architecture consisting of two processes the MIC is defined as the second order difference of the mean response time under manipulation of the speed of each process. Formally, we are interested in two random variables representing the duration of the two mental processes (T_x, T_y) and the random variable representing the time to respond with both sources of information (T_{xy}). The duration of each process is manipulated through the manipulation of the two external external factors (f_x and f_y) that are binary valued (Low and High). It is experimentally provided that the Low factor value leads to slower response time then the High factor value.

$$\Delta^2 \mathrm{E}\left[T_{xy}; f_x, f_y\right] = (\mathrm{E}\left[T_{xy}; f_x = \mathrm{Low}, f_y = \mathrm{Low}\right] - \mathrm{E}\left[T_{xy}; f_x = \mathrm{Low}, f_y = \mathrm{High}\right]) - (\mathrm{E}\left[T_{xy}; f_x = \mathrm{High}, f_y = \mathrm{Low}\right] - \mathrm{E}\left[T_{xy}; f_x = \mathrm{High}, f_y = \mathrm{High}\right])$$

For more practical purpose, the above equation could be written in the form of mean response times, where RT indicates the mean observed response time and the subscript indicate the factor levels for f_x and f_y,

$$\mathrm{MIC} = (\mathrm{RT}_{\mathrm{LL}} - \mathrm{RT}_{\mathrm{LH}}) - (\mathrm{RT}_{\mathrm{HL}} - \mathrm{RT}_{\mathrm{HH}}) = \mathrm{RT}_{\mathrm{LL}} - \mathrm{RT}_{\mathrm{LH}} - \mathrm{RT}_{\mathrm{HL}} + \mathrm{RT}_{\mathrm{HH}}.$$

To draw meaningful inferences based on the MIC, an important condition of selective influence must hold (Townsend & Thomas, 1994; Dzhafarov, 2003; Dzhafarov et al., 2004). In general, the condition of effective selective influence requires that a single external factor exclusively affects only one subprocess and that affect has some measurable consequence. In our two-process mental architecture example, effective selective influence means that by varying between the two values (Low and High) the experimental manipulations f_x and f_y exclusively affect only their respective processing times T_x and T_y. Further, the manipulation must have an affect, i.e., $(T_x; f_x = \mathrm{Low}) > (T_x; f_x = \mathrm{High})$. The difference in processing time (for each T_x, T_y) between low and H levels of manipulation is in the relevant literature is referred to as the saliency effect. In some cases, selective influence can be directly assessed (Dzhafarov & Kujala, 2010, 2014), although in the general case, it is only possible to test for violations of the condition. The most common approach is to check for stochastic dominance between $(T_{xy}; f_x = \mathrm{Low}, f_y = \mathrm{Low})$ and T_{xy} when either $f_x = \mathrm{High}$ or $f_y = \mathrm{High}$:

$$P(\mathrm{RT}_{\mathrm{LL}} \leq t) \geq P(\mathrm{RT}_{\mathrm{LH}} \leq t)$$
$$P(\mathrm{RT}_{\mathrm{LL}} \leq t) \geq P(\mathrm{RT}_{\mathrm{HL}} \leq t)$$

as well as between T_{xy} when either $f_x = \mathrm{Low}$ or $f_y = \mathrm{Low}$ and $(T_{xy}; f_x = \mathrm{High}, f_y = \mathrm{High})$,

$$P(\mathrm{RT}_{\mathrm{LH}} \leq t) \geq P(\mathrm{RT}_{\mathrm{HH}} \leq t)$$
$$P(\mathrm{RT}_{\mathrm{HL}} \leq t) \geq P(\mathrm{RT}_{\mathrm{HH}} \leq t).$$

This condition is implied by selective influence(Townsend & Thomas, 1994; Schweickert et al., 2000)[1]

The sign of the MIC is used to diagnose two of the fundamental properties in cognitive operations. When each subprocess is selectively influenced in a serial system, then the MIC will be zero (regardless of stopping rule), whereas in a parallel system the MIC will be non-zero. Parallel, exhaustive processing leads to $\mathrm{MIC} < 0$ and parallel, first-terminating processing leads to $\mathrm{MIC} > 0$. Like the parallel, first-terminating processes, coactive processes will also lead to $\mathrm{MIC} > 0$.

[1] See Heathcote et al. (2010) for a survey of approaches to testing stochastic dominance.

SFT includes a more powerful statistic to diagnose processes, the survivor interaction contrast function, SIC(t). The SIC can be estimated from the empirical survivor (or conversely 1− empirical survivor = empirical cumulative distribution function),

$$\hat{S}(t) = \frac{\#RT > t}{\#RT} = 1 - \frac{\#RT \leq t}{\#RT} = 1 - \hat{F}(t).$$

In which # stands for number of response trials observed. To calculate the empirical SIC, empirical survivor functions are calculated for each factorial condition, and used in the form of the second order difference analogously to the MIC (Houpt & Townsend, 2010; see also Houpt et al., 2017, 2016 for Bayesian alternatives).

$$\mathrm{SIC}(t) = \Delta^2 \mathrm{S_{xy}}\,(\mathrm{t}) = [\mathrm{S_{LL}}(\mathrm{t}) - \mathrm{S_{LH}}(\mathrm{t})] - [\mathrm{S_{HL}}(\mathrm{t}) - \mathrm{S_{HH}}(\mathrm{t})]$$

The subscript indicate the factor levels for f_x and f_y (Low and High) associated with each subprocess of interest. The relationship between the SIC and MIC is straightforward, MIC = $\int_0^\infty \mathrm{SIC}(t)\ dt$. This relationship makes it clear that the SIC provides at least as much information as the MIC. Indeed, unlike the MIC described above, all five canonical mental architectures could be distinguished based on the shape of SIC function. For example, serial exhaustive and serial first-terminating function, both predict MIC = 0, but predict different SIC functions.

While the SIC has more diagnostic power, the MIC has some advantages over the SIC for diagnosing underlying mental architectures. First, fewer trials are needed to achieve a good estimate of the MIC because it is a single value, unlike the SIC which is an entire function. In practice this means that running a study using MIC could require fewer trials then a study using SIC. If there is little constraint on the number of trials that can be collected, SIC might be preferred (e.g., Townsend & Fifić, 2004). In many cases, conducting a large scale study involving a large number of stimulus trials per subject is not a realistic scenario. Research participants, are usually reluctant to participate in lengthy studies, and are more likely to drop out. Hence, long term studies can require significant financial compensation to recruit and retain participants. Additionally, subjects from particular populations are only available for study participation for a brief period of time. This can be due to limited mental capabilities and are not able to focus for a long period of time, or due to other constraints on their time. For example, autistic children (cf. Johnson et al., 2010), or air force pilots (cf. Schreiber et al., 2006) would only be available to serve as experimental participants for a limited number of trials. In such situations it is Hly impractical to conduct repeated study sessions limiting a researcher to a relatively smaller number of response trials .

1.1. Existing approaches to statistical inference with the SIC and MIC

A number of approaches have been introduced for making inferences based on the SIC and MIC (see Houpt & Burns, 2017, for a review). The initial approach

to testing the MIC values relied on using a factorial ANOVA design. ANOVA is an almost natural choice given that the factorial nature of an SFT study's manipulations. ANOVA is used to test the hypothesis on whether or not an observed MIC value significantly departs from zero value, which was identified as the null-hypothesis (cf. Kirk, 2012). An alternative, nonparametric approach was to use bootstrapping (see Van Zandt, 2002, for details) to construct confidence intervals around observed MIC values. If zero is within the confidence intervals of the estimated MIC, a researcher would fail to reject null-hypothesis, otherwise the null is rejected and the sign of the MIC value determines whether the MIC shows overadditivity, or underadditivity (see, e.g., Yang et al., 2014a, 2012). An alternative, nonparametric test, based on a generalization of the Kolmogorov-Smirnov test, has also been proposed as an approach to analyzing the SIC shape, and hence whether the MIC is significantly different from 0 (Houpt & Townsend, 2010). Houpt & Townsend (2010) also compared standard ANOVA and nonparametric interaction tests for testing the null-hypothesis that $\mathrm{MIC} = 0$.

There are two main limitations of these existing approaches. The first limitation is related to the statistical inference and the diagnostic power of the SFT nonparametric methods. Although very useful at the initial stages of the development of the SFT technology, statistical inference based on null-hypothesis testing can be limiting. Using the ANOVA and bootstrapping approaches described above the null-hypothesis is exclusively linked to one mental architecture $\mathrm{MIC} = 0$, which is the signature of serial processing. A significant result would indicate that processing is not serial, but there is no way to reject parallel processing: A classical failure to rejecting the null hypothesis, that is likelihood that $\mathrm{MIC} = 0$ given the null is true, doesn't imply that that the alternative hypothesis is not true $\mathrm{MIC} \neq 0$, given the data.

To address such Bayesian arguments other alternative analyses have recently been proposed for the SIC. Houpt et al. (2016) proposed a semiparametric Bayesian approach for estimating posterior distributions over SICs. Houpt et al. (2017) have also developed parametric and nonparametric Bayesian approaches to estimating SIC shape. However, neither of these approaches fully address the second limitation, which is the inability of the current methods to make group level inferences that involves quantitative statistical description of a sample, that can be used to generalize to the population.

Until recently the SFT approach has been focused on individual subject analysis, in addition to the statistical inference about the underlying cognitive operations. Indeed, the many SFT studies made a final conclusion in the form of basic descriptive statistics, nominally classifying subjects based on their achievement. For example, a short-term memory study indicated individual differences across and within experimental conditions of different short-term memory manipulations. The major finding was that some subjects would switch from serial to parallel when the timing condition was changed (Townsend & Fifić, 2004). Although these results are very useful, the nominal categorization based on the statistical inferences using the null-hypothesis test, tells little about the population from which the subjects had been sampled.

To summarize, the two limitations of statistical inference with SIC/MICs have been discussed, the first one being logically limited commitment to the null-hypothesis testing, and the second one being the lack of group level analysis. Both limitations can jeopardize the practical power of the SFT method, with possibility to systematically biasing inferences.

To address these limitations, we propose hierarchical Bayesian analysis. Hierarchical modeling allows for compromise between modeling individual differences and group level information (cf. Busemeyer & Diederich, 2010, Chapter 6). By employing a Bayesian approach, we can use priors to incorporate information about task constraints on a likelihood that some fundamental processing property are present. For example, when exhaustive processing is required by the task and accuracy is H, there is low prior probability that first-terminating processing was employed, hence MIC is less likely to be positive. A Bayesian approach can be used estimate posterior probabilities of each category of MIC (less than, greater than or equal to zero) rather than being limited to testing the null-hypothesis that MIC = 0. These MIC posterior probabilities can be estimated at both the individual level, indicating how likely each MIC category is for each subject, as well as the group level.

In the next section, we will describe the hierarchical Bayesian model for the MIC, then we will examine the modeling approach applied to simulated data and a dataset that is commonly used to validate SIC statistics.

2. Hierarchical Bayesian MIC

Our full model is given in Table 1 and depicted in Figure 1. The central component of the model is a linear model of the mean response time, much like an ANOVA (cf. Rouder et al., 2012). We derived this linear model based on two principles. First, the MIC is the main variable of interest, so we needed it to be explicitly represented. This allows us to set priors on both its category and magnitude. Second, we ensured that the variability of the prior on the mean for each condition would not be different across the salience levels. There are a number of different possibilities for this matrix. For our purposes, we chose,

$$\begin{pmatrix} \text{MIC} \\ \Delta_2 \\ \Delta_1 \\ \mu \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1/2 & 1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \begin{pmatrix} \mu_{HH} \\ \mu_{HL} \\ \mu_{LH} \\ \mu_{LL} \end{pmatrix}.$$

Here, MIC is the mean interaction contrast, Δ_1 is the average increase in mean response time due to a change in salience on process 1 across salience levels on Channel 2 (and likewise for Δ_2), and μ is the grand mean response time.

Thus, if we set our priors on the MIC, Δ_i and grand mean, they can be translated into priors for the mean RT at each salience level using the inverse

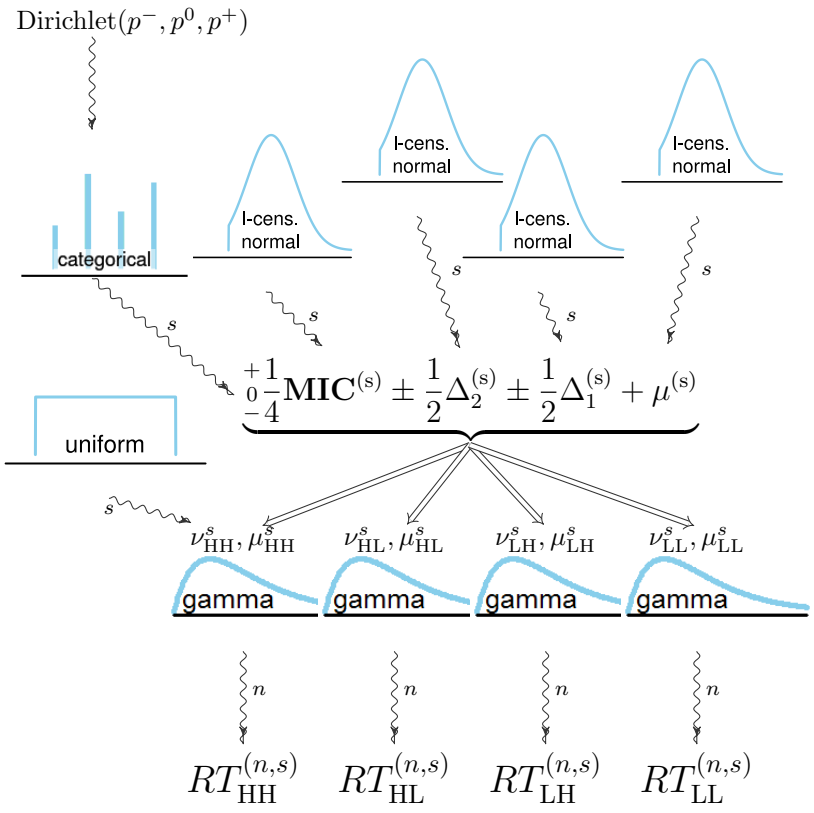

Figure 1: Diagram indicating the hierarchical structure of the Bayesian model of RT from which we deduce information about the MIC. Wiggling lines indicate a random relationship (e.g,. RTs are sampled from a gamma distribution), and double, straight arrows indicate a deterministic relationship (e.g., μ_{HH} is determined by the linear model at the center of the diagram). Note that there is a separate prior for the rate multiplier of each RT gamma distributions, although only one is depicted to reduce clutter. Thanks to John Kruschke (http://doingbayesiandataanalysis.blogspot.com/), Rasmus Bååth (http://www.sumsar.net/about.html) and Tinu Schneider (https://github.com/tinu-schneider/DBDA_hierach_diagram) for making this figure possible.

of the mapping above,

$$\begin{pmatrix} \mu_{HH} \\ \mu_{HL} \\ \mu_{LH} \\ \mu_{LL} \end{pmatrix} = \begin{pmatrix} 1/4 & -1/2 & -1/2 & 1 \\ -1/4 & 1/2 & -1/2 & 1 \\ -1/4 & -1/2 & 1/2 & 1 \\ 1/4 & 1/2 & 1/2 & 1 \end{pmatrix} \begin{pmatrix} \text{MIC} \\ \Delta_2 \\ \Delta_1 \\ \mu \end{pmatrix}.$$

For the MIC, we set up the likelihood as a mixture model of three categories (χ) for each subject (superscript-s), one in which the MIC > 0, one with MIC < 0 and one with MIC $= 0$. Each subject's data had its own categorical distribution over the three cases with a Dirichlet prior over the case probabilities, ($\{p^-, p^0, p^+\}$). Each of those Dirichlet priors were drawn from a single Dirichlet distribution representing the group level. The main idea of this approach is to consider each category a potentially believable structure, then use the mechanisms of MCMC to estimate the posterior probability associated with each category (cf. Kruschke, 2010, Chapter 12).[2]

Assuming the RTs are on a millisecond scale, the prior on the magnitude of the MIC in the two cases for which it was non-zero, was a truncated Gaussian with mean 100 and standard deviation 50. Although a separate random variable was used for the MIC magnitude depending on whether it was for the positive case or negative case, all three three cases shared the same Δ and μ parameters. The priors on Δ_1 and Δ_2 each had a truncated Gaussian distribution with the same parameters (mean 100, standard deviation 50). For the grand mean μ, we use a truncated normal distribution with mean 400 and standard deviation 100.

In theory, the mean response time for a particular subject in a condition could be negative under this model, e.g., when the effect of the salience manipulations has Her magnitude than the grand mean response time. Although this possibility should have no probability in the data, it is important to constrain the parameters of the prior distributions so that a negative mean response time is unlikely or impossible.

For the likelihood of the response times, we used a gamma distribution, which has the skewed shape commonly observer in RTs and only has support on positive values. The gamma distribution has two parameters, usually a shape and a rate, hence one additional parameter was required. In this case, we chose to use the rate as the additional free parameter. The standard rate/shape parameterization of the gamma distribution could then be recovered because the mean of a gamma distribution is the shape parameter divided by the rate parameter. Like the Δ parameters, we used only a single rate multiplier across the three MIC cases. For the analyses reported below, we chose improper flat priors over the positive real line for the rate multipliers to allow flexibility in how the shape and rate traded off for a given mean RT.

[2] JAGS and BUGS allow one to specify categorical priors directly, however due to the sampling mechanism, it is not possible in Stan. To implement mixture models in Stan, one can marginalize over the categorical parameter, leaving the category probability parameters to remain without a variable explicitly representing the category. See Stan Development Team, 2015, Section 10 for details.

$$
\begin{aligned}
\left\{p^{-}, p^{0}, p^{+}\right\} &\sim \text{Dirichlet}\left(0.25, 0.5, 0.25\right) \\
\left\{p^{-,s}, p^{0,s}, p^{+,s}\right\} &\sim \text{Dirichlet}\left(p^{-}, p^{0}, p^{+}\right) \\
\chi^{s} &\sim \text{categorical}\left(p^{-,s}, p^{0,s}, p^{+,s}\right) \\
\text{MIC}^{s} &\sim \text{truncated normal}(100, 50) \\
\Delta_2^{s} &\sim \text{normal}(-100, 50) \\
\Delta_1^{s} &\sim \text{normal}(100, 50) \\
\mu^{s} &\sim \text{normal}(200, 25) \\
\nu_{x,y}^{\chi,s} &\sim \text{gamma}(1, 1) \\
\text{RT}_{\text{x,y}}^{(\text{n,s})} &\sim \text{gamma}(\nu_{x,y}^{\chi,s}, \mu_{x,y}^{\chi,s})
\end{aligned}
$$

$$
\begin{aligned}
\mu_{x,y}^{+,s} &= (-1)^{\{x \text{ is H}\}}(-1)^{y \text{ is H}}\frac{1}{4}\text{MIC}^{s} + (-1)^{\{y \text{ is H}\}}\frac{1}{2}\Delta_2^{s} + (-1)^{\{x \text{ is H}\}}\frac{1}{2}\Delta_2^{s}\Delta_1^{s} + \mu^{s} \\
\mu_{x,y}^{-,s} &= (-1)(-1)^{\{x \text{ is H}\}}(-1)^{y \text{ is H}}\frac{1}{4}\text{MIC}^{s} + (-1)^{\{y \text{ is H}\}}\frac{1}{2}\Delta_2^{s} + (-1)^{\{x \text{ is H}\}}\frac{1}{2}\Delta_2^{s}\Delta_1^{s} + \mu^{s} \\
\mu_{x,y}^{0,s} &= b_{xy}\Delta_2^{s} + c_{xy}\Delta_1^{s} + \mu^{s}
\end{aligned}
$$

Table 1: Complete description of the model. Above, the prior distribution for each parameter is listed. Below, the formula for the mean of the response time distributions as a function of the the parameters is given.

3. Application to Simulated Data

To better understand how well this model can be used to assess MIC category, and hence discriminate serial and parallel processing, we tested it on a series of simulated data. We varied the architecture and stopping rule for processing two sources of information, the parameters of interest that determine the MIC category. Recall that selectively influenced serial models imply MIC = 0 regardless of stopping rule, parallel models with exhaustive stopping rules imply MIC < 0 and parallel models with first-terminating rules imply MIC > 0. In addition to the sign of the MIC, other parameters can influence the magnitude of the MIC and precision with which it can be measured. One of the most important parameters is the effectiveness of the salience manipulation, i.e., how much faster each source of information is processed in a H salience condition relative to a low-salience condition. Additionally, the amount of data, particularly the number of response times collected from each subjects and the total number of subjects was varied.

3.1. Method

Data were generated assuming either 10, 15, or 20 subjects. For each simulated subject, either 40, 50, 60, or 70 response times were simulated per condition (e.g., 70 in the HH condition, 70 in the HL condition, 70 in the LH condition, and 70 in the LL condition). Each response time was simulated by combining the subprocess durations (T_1, T_2) according the corresponding architecture and stopping rule:

Parallel, Exhaustive:	$\mathrm{RT} = \max(T_1, T_2)$
Parallel, First-Terminating:	$\mathrm{RT} = \min(T_1, T_2)$
Serial, Exhaustive:	$\mathrm{RT} = T_1 + T_2$
Serial, First-Terminating:	$\mathrm{RT} = T_1$ with probability .5; $\mathrm{RT} = T_2$ otherwise

Within each dataset, all simulated subjects had the same architecture and stopping rule.

Subprocess durations were generated assuming the completion times were based on the first passage time of a Brownian motion process, and hence followed an inverse Gaussian distribution,

$$f(t; \alpha, \nu) = \frac{\alpha}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[\frac{-(\alpha - \nu t)^2}{2t\sigma^2}\right].$$

The threshold activation for a response, α, was set to 30 and the diffusion coefficient, $\sigma^2 = 1$ for all simulations. The drift rate, ν, depended on the condition. To simulate a low salience trial for a subprocess, the drift rate was

set to 0.1. For H salience trials, the drift rate was set to either $1.5, 2, 2.5$, or 3 times the low salience drift rate.[3]

The Bayesian analyses were run using Stan (Stan Development Team, 2015, 2014) on a combination of MindModeling.org (Harris, 2008), the Oakley cluster at the Ohio Supercomputing Center (Ohio Supercomputer Center, 2012, 1987), and Microsoft's Azure service.[4] Follow-up analyses were done using R statistical software (R Development Core Team, 2011) and the sft R package (Houpt et al., 2013). The Stan code is included as supplementary material. We ran four chains using $10,000$ warm-up samples and $20,000$ additional iterations per chain.[5] All chains were visually assessed for mixing and Gelman-Rubin $\hat{R}$ values were less than 1.01 for all parameters.

3.2. Results

A summary of the group level posterior and subject level posterior are shown in Figure 2 and Figure 3 respectively. Each row corresponds to a different model used to generate the data. The left column gives the mean posterior probability that the MIC is in the category predicted by the generating model (e.g., $MIC > 0$ for data generated from a parallel first-terminating model). The right column indicates the standard deviation of the posterior probability of that MIC category. In the subject level data, the values are averaged across the simulated subjects (i.e., the mean posterior probability is the average across subjects of their individual mean posterior probability; the standard deviation is the average across subjects of the standard deviation of the posterior probability that their MIC is in the given category).

The only parameter that had a clear effect on the posterior probability over MIC category, for both the group and individual level, was the strength of the salience manipulations (indicated by line darkness in Figures 2 and 3). At the lowest manipulations strength, the most likely MIC is 0 for all of the models, regardless of the number of subjects or the number of trials per distribution. The posterior probability correct of positive and negative MICs increases essentially linearly with an increase in salience for the parallel–exhaustive data and parallel–first-terminating data respectively. In the serial, first-terminating data, the posterior probability stays essentially flat between 0.6 and 0.8 for the range of salience. Interestingly, there seems to be a negative trend in the serial-exhaustive data, particularly with only 50 trials per distribution.

[3] Increasing the drift rate while holding the threshold constant produces stochastic dominance for this model. With $\Phi()$ indicating the standard normal CDF, the CDF of the first passage time is

$$F(t;\alpha,\nu) = \Phi\left[t^{-1/2}(t\nu-\alpha)\right] + e^{2\alpha\nu}\Phi\left[t^{-1/2}(t\nu+\alpha)\right].$$

Φ and exp are both monotonically increasing functions and increasing ν increases the argument of each term and hence F.

[4] `http://azure.microsoft.com`

[5] For more details on the parameters of a Stan analysis, see Stan Development Team (2015).

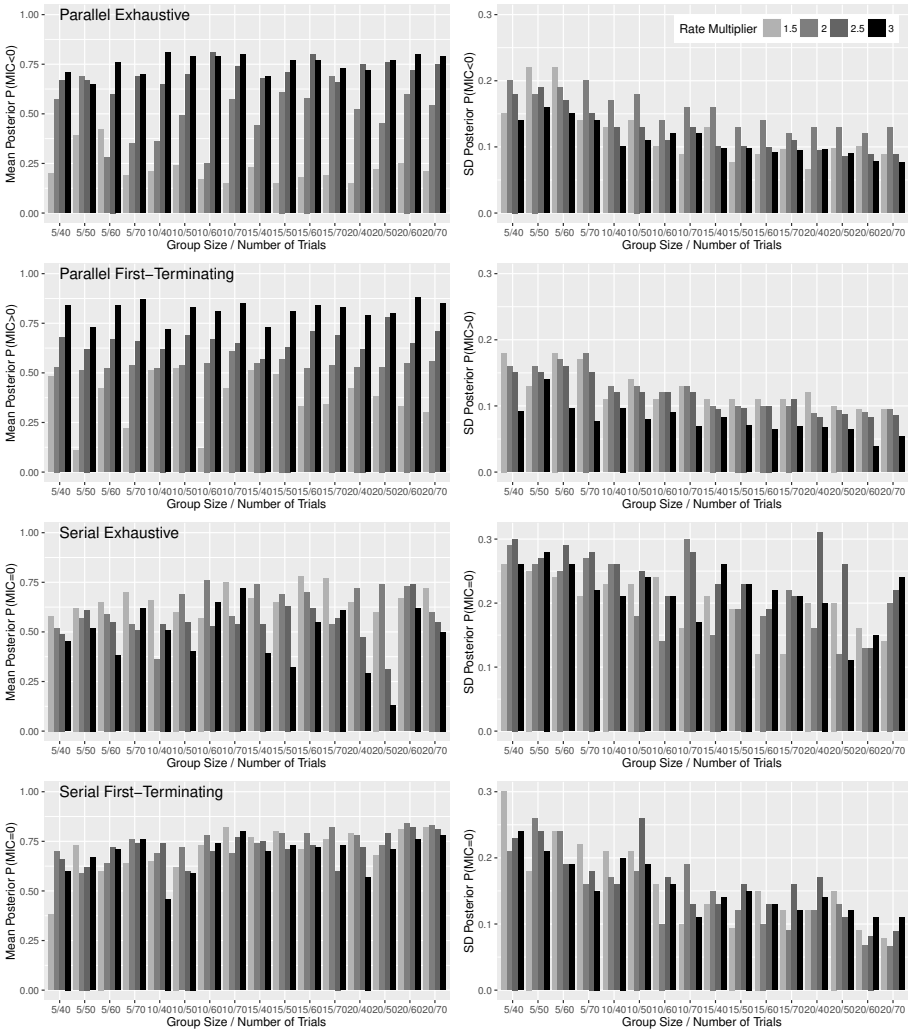

Figure 2: Simulation results for the group level probabilities. Each row corresponds to a model that was used to generate the data. The left column shows the mean posterior probability that the group MIC is in the category predicted by the model that generated the data. The right column shows the standard deviation of the that posterior probability. Within each panel, bars are grouped by the number of trials per subject, then by the number of subjects per group. The rate multiplier, representing the strength of the salience manipulation, is indicated by the shade of the bars.

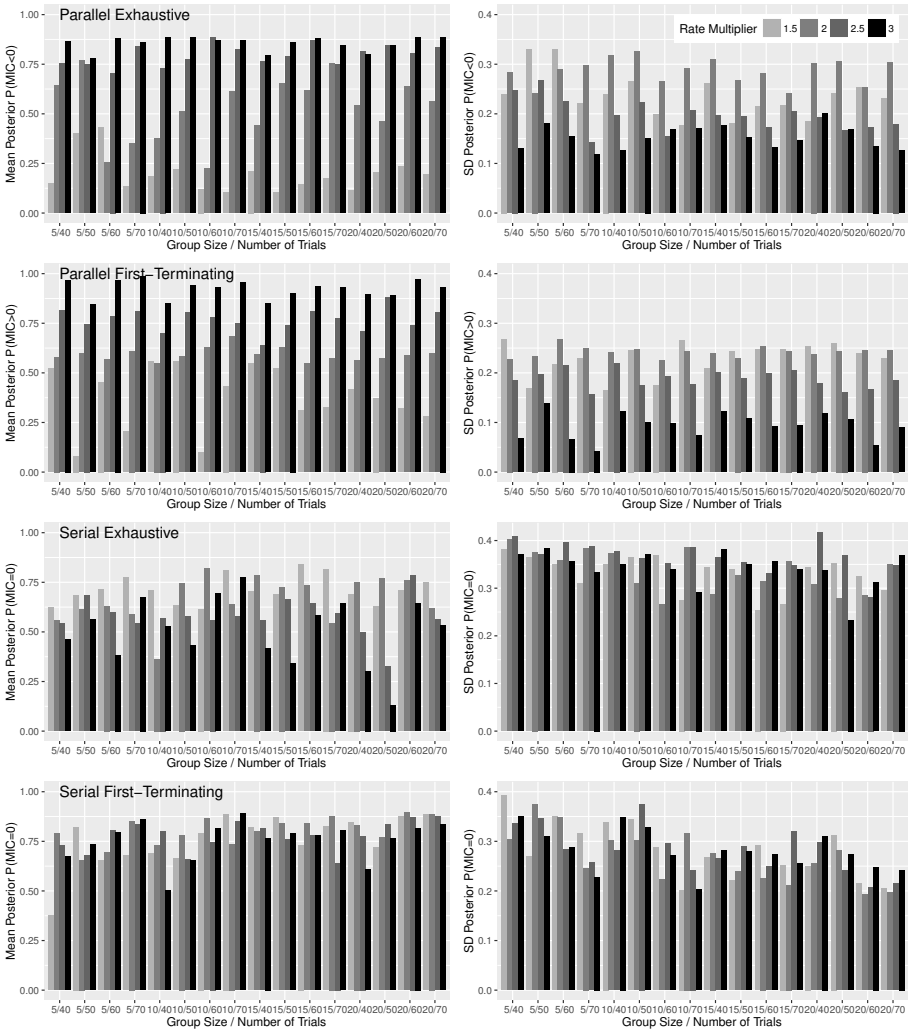

Figure 3: Simulation results for subject level probabilities. As in the previous figure, row corresponds to generating model. In this figure, the left column shows the posterior probability that the subject MIC is in the category predicted by the model that generated the data, averaged across subjects. The right column shows the standard deviation of the that posterior probability averaged across subjects. Within each panel, bars are grouped by the number of trials per subject, then by the number of subjects per group. The rate multiplier, representing the strength of the salience manipulation, is indicated by the shade of the bars.

The standard deviation of the category probability was also affected by the salience strength. In the parallel model data, the lowest rate multiplier resulted in lower standard deviations, reflecting more certainty in the posterior that the MIC was zero. This is likely due to the fact that the differences $RT_{LL} - RT_{LH}$ and $RT_{HL} - RT_{HH}$ in the data are not large enough to make their differences (the interaction) detectably different from zero. For the larger rate multipliers in the parallel data, the standard deviation was again smaller, but in this case reflecting more certainty that the MIC was negative or positive for the exhaustive and first-terminating data respectively.

In addition to the rate multiplier, the number of subjects affected the group level and the number of trials per subject affected the subject level. More subjects led to lower standard deviations at the group level, and lower standard deviations at the subject level, although the effect was more prominent at the group level. More trials per subject led to lower standard deviations at the subject level, but had little affect at the group level.

In general, we find these results quite promising. Most experiments relying on SICs use at 100 or more trials per distribution and approximately 10 subjects (e.g., Yang et al., 2011). Our results indicate that, as long as the salience manipulation is sufficient, this is enough data for drawing both group and individual level inferences. The results regarding the rate multiplier indicate an important cautionary note as well. If the salience manipulation is not strong enough, data from any of the four generating models will be classified as having a zero MIC. Hence it is important to aim for strong salience manipulations in designing experiments to be analyzed with this (or any other SIC) analysis. Based on the impact of the rate multiplier, when the salience is strong, the model should do well.

4. Application to data from a simple detection experiment

One of the standard data sets for testing SFT statistical analyses is the dot detection data reported in Eidels et al. (2015, Study I), which is available in the sft R package (Houpt et al., 2014; R Development Core Team, 2011). In this study, one or two small, low-contrast dots were shown on a uniform background either above the mid-line of the display, below the mid-line, or both. Each dot could be displayed at a slightly Her contrast (H salience) or lower contrast (low salience). There were three factors manipulated within subjects: dot presence (present, absent); dot salience (H, low); and task instructions (OR and AND). The task instructions were held constant within a day. For example, on one day participants were asked to respond affirmatively if either dots was shown and negatively otherwise (OR rule). On another day, participants were asked to respond affirmatively only if both dots were shown and negatively otherwise (AND rule).

The simple detection study allows for the model assessment by inspecting the observed MIC values. If the participants were processing the visual stimuli in parallel, we would expect a positive MIC in the OR condition and a negative MIC in the AND condition. It is also possible that despite the "OR"

	AND Task			OR Task		
	+	0	−	+	0	−
Group	0.14	0.06	0.80	0.73	0.11	0.17
S1	0.04	0.00	0.95	0.93	0.01	0.06
S2	0.04	0.06	0.90	0.53	0.38	0.09
S3	0.06	0.02	0.92	0.94	0.01	0.05
S4	0.02	0.02	0.95	0.63	0.27	0.10
S5	0.02	0.02	0.97	0.94	0.04	0.02
S6	0.06	0.11	0.83	0.95	0.03	0.02
S7	0.03	0.04	0.93	0.78	0.15	0.07
S8	0.05	0.08	0.87	0.92	0.05	0.03
S9	0.06	0.01	0.94	0.93	0.03	0.04

Table 2: Mean posterior probabilities of MIC category when AND and OR conditions were analyzed separately. Note that although category probabilities must sum to one for each posterior sample in the MCMC chain, rounding error means that these posterior mean probabilities may sum to values slightly different from one.

instructions, the participants used an exhaustive stopping rule in that condition, in which case we would expect a negative MIC in both conditions. In the AND condition, the participants would have low accuracy if they used a first-terminating stopping rule, which was not the case. However, if participants were using a coactive strategy, then a positive MIC would be indicated in the AND condition. If a participant used a serial strategy, either exhaustive or first-terminating, the resulting MIC would be 0. For estimating the MIC, there were 200 trials for each condition of interest (HH, HL, LH, and LL) for each instruction type. The data set provided results that are consistent across subjects, and clearly identifiable using the SFT approach. As such the data set provides a valuable validation tool for the new analysis.

In our initial application of the new hierarchical analysis to the Eidels et al. (2015) data, we separately analyzed the AND condition and the OR condition. As in the simulations section, we ran four chains using $10,000$ warm-up samples and $20,000$ additional iterations per chain. All chains were visually assessed for mixing and Gelman-Rubin $\hat{R}$ values were less than 1.01 for all parameters.

Results of the first analysis are reported in Table 2 and are consistent with previous analyses based on non-Bayesian methods (Houpt & Townsend, 2010; Houpt et al., 2016). For the AND task, the posterior probabilities strongly favored the negative MIC at the group level and for each of the individuals. Similarly, for the OR task, positive MICs had the Hest probability at the group level and for each of the individuals. Two participants, S2 and S4, had relatively lower probabilities of positive MICs in the OR task, with posterior odds ratios of 2.8 and 10 respectively for positive over zero MICs. On the whole, there is strong evidence against serial processing (which implies MIC = 0). Further, there is even stronger evidence against coactive processing in the AND task ($MIC > 0$) or exhaustive processing in the OR task (MIC < 0).

Given that the model indicated the same MIC category across participants,

	+	0	−
Group	0.46	0.06	0.48
S1-AND	0.04	0.00	0.95
S2-AND	0.09	0.05	0.86
S3-AND	0.09	0.02	0.90
S4-AND	0.06	0.02	0.93
S5-AND	0.04	0.01	0.94
S6-AND	0.15	0.11	0.74
S7-AND	0.07	0.04	0.89
S8-AND	0.12	0.07	0.81
S9-AND	0.07	0.01	0.92
S1-OR	0.93	0.01	0.06
S2-OR	0.56	0.28	0.17
S3-OR	0.94	0.01	0.05
S4-OR	0.58	0.17	0.25
S5-OR	0.93	0.02	0.05
S6-OR	0.93	0.02	0.05
S7-OR	0.77	0.08	0.15
S8-OR	0.91	0.02	0.07
S9-OR	0.93	0.01	0.06

Table 3: Mean posterior probabilities of MIC category when AND and OR conditions were analyzed as a samples from the same group distribution. S*i*-AND indicates data from the AND instructions while S*i*-OR indicates data from the OR instructions. The model did not encode the relationship between AND and OR data from the same subject. Note that although category probabilities must sum to one for each posterior sample in the MCMC chain, rounding error means that these posterior mean probabilities may sum to values slightly different from one.

one may wonder whether the hierarchical model is biased toward assuming a single MIC category for all participants. While a bias toward homogeneity could be intentionally built into the model by using a group level prior with most of the probability mass focused on a particular MIC category, the prior we used was meant to allow variability across subjects. To explore the possibility that the model is biased toward homogeneity, we recoded the Eidels et al. (2015) data so that each participant–instruction combination was treated as a separate member of a single group. I.e., the data from Subject 1 in the OR condition was recoded as S1-OR while the data from him/her in the AND condition was recoded as S1-AND, and likewise for the other 8 participants.[6] We ran four chains using $10,000$ warm-up samples and $20,000$ additional iterations per chain. All chains were visually assessed for mixing and Gelman-Rubin $\hat{R}$ values were less than 1.01 for all parameters.

[6]Although we could have built structure into the model relating a subject's performance across the instructions, we chose to treat the RTs for a given subject with a given instruction as conditionally independent given the group MIC value.

The posterior probabilities in Table 3 indicate very little probability of a zero MIC, but roughly equal probabilities of positive and negative MICs at the group level. This is noteworthy for two reasons: First, it demonstrates that the model does not inherently predict homogeneity. Second, it illustrates the advantage of using a categorical prior for the sign of the MIC because the positive and negative individual MICs were not averaged (which would give a group MIC near zero). Despite the fact that the posterior probabilities indicate heterogeneity, there was still some shrinkage in the individual posterior probabilities: For the AND data, the probability of a negative MIC was slightly smaller and slightly larger for positive MICs while the opposite was true for the OR data. The probability of a zero MIC stayed was roughly the same for the AND data as in Table 2. The probability of a zero MIC in the OR data decreased some, particularly for those participants for who had slightly Her posterior probabilities of a zero MIC on the OR condition in Table 2. It is clear that this model does not impose homogeneity on the individuals.

On the whole, these results are quite promising. The model was able to estimate a reasonable group level and individual level posterior distribution. These results provide converging evidence with the previously reported analyses of these data, which had shown parallel processing for all participants and the appropriate stopping rule application for the specific stopping rule task instruction condition. The additional benefit of the new Bayesian hierarchical approach is that it provides not only the individual level information, but also the group level information.

5. Discussion

The survivor and mean interaction contrasts are among the most powerful diagnostic methods for discerning whether people using information in parallel or in series because they avoid the model mimicking dilemma that plagues other methods. However, the interaction contrast approach complicates the statistical analysis so methods for statistical inference have been relatively lacking until recently. Houpt & Townsend (2010) proposed a null-hypothesis test for the SIC and compared ANOVAs and adjusted-rank-transform tests for the MIC. More recently Houpt et al. (2017) and Houpt et al. (2016) proposed Bayesian analyses, but all of these approaches are for only individual level analysis.

In this paper we have addressed one of the major outstanding issues in the statistical analysis of MICs, the lack of an approach to make group level inferences. We demonstrated the efficacy of a hierarchical Bayesian model of the MIC for making both individual level and group level inferences with a relatively small number of trials and subjects, using both a simulation study and an application to a standard dataset. Performance of the analysis on the simulated data improves with having more subjects, trials, and increased efficacy of the salience manipulation. Nonetheless, with just 50 trials per condition, inferences based on the model's posterior probability of of the MIC associated with the data generating process led to quite satisfactory results.

Both the SIC and MIC measures are frequently used as an individual subject assessment to indicate qualitative differences in cognitive operations in a sample of subjects. An obstacle in assessment of individual human subjects' cognitive operations is the requirement for a large number of trials per subject. For example, Houpt & Townsend (2010) demonstrated their statistical analysis with 200 trials per distribution, which when trials are balanced appropriately (cf. Houpt & Townsend, 2012; Mordkoff & Yantis, 1991) can mean 3200 trials per participant. While this samples size would not cause a psychophysicist to balk, many interesting populations, such as clinical groups, experts, and some age groups, are available only for a limited time, and thus permit only a smaller set of observations per individual. Although it has less diagnostic power then SIC, the MIC can rely on a small data sets, making it a more practical measure for cases in which only limited numbers of trials are available per subject.

One unexpected finding was that with increased salience manipulation efficacy but a limited number of trials, MIC category recovery performance weakened for the data generated from a serial exhaustive process. As the stimulus salience effect increased, the posterior probability of a zero MIC decreased. The extent to which this is a property of the particular assumptions we have made, either in generating the data or the model itself, or if it is an outcome specific to this sample dataset, will be an interesting topic of further investigation.

In addition to the simulated data, the model performed well on the SFT data that is commonly used to assess SIC and MIC statistics from Eidels et al. (2015). The Bayesian hierarchical MIC model exhibited strong convergence to the conclusions drawn from SIC level analysis in other papers (Houpt & Townsend, 2012; Houpt et al., 2017, 2016). Perhaps the most challenging test of the model was its application to heterogeneous experimental conditions in which the subjects were using different processes. In the Eidels et al. (2015) study, two experimental conditions were imposed by the instructions. In the OR condition, subjects could use a first-terminating stopping rule, while in the AND condition they have to use an exhaustive stopping rule. To test whether the model is able to detect variation across subjects, the data in each condition were treated as coming from the same group, thus having heterogeneous subject properties. When the hierarchical Bayesian MIC model was applied to the data in this format, the analysis appropriately identified the expected MIC category at the subject level and indicated approximately 50% posterior probability for each of the positive and negative MIC categories at the group level. This demonstrated that the Bayesian MIC model can identify individual subjects' differences within a group data set, and will not always indicate that all subjects use the same cognitive operations.

Our approach to exploring the individual and group level MIC analysis using the hierarchical Bayesian MIC model is similar in many ways to the method proposed by (Thiele & Rouder, 2016).[7] The overarching goals of both approaches

[7]Both research groups independently developed research approaches to extending the SFT MIC tests using the hierarchical Bayesian model, and discovered each others work through

are the same: 1) To better quantify the evidence for either serial or parallel processing at the group level 2) Rein in the bias toward heterogeneity that results from analyzing subjects as unrelated. Similarly, the structure of the models are quite similar, with a linear model predicting the mean processing time across distributions within a subject. The main distinction between the two approaches is the focus on determining whether architecture is homogeneous or heterogeneous across participants (Thiele, 2014) and estimating posterior probabilities associated with each model (herein) when heterogeneity is given positive prior probability (herein). There are also some minor differences between the two models. First, Thiele & Rouder (2016) use a normal distribution as their model of the response times where as we use a gamma distribution. They report choosing the normal distribution for two reasons, computational tractability and the ease with which the sign of the MIC can be constrained relative to non-normal distributions. From our perspective, the computational power of Stan and Hamiltonian Monte-Carlo methods means that we can use a more realistic distribution for response times and still obtain results from the analysis in a reasonable time frame. Furthermore, our categorical approach using the Dirichlet prior allows us to model the sign of the MIC without additional difficulty in implementation. The second difference between the approaches is the means by which conclusions are drawn. The Thiele & Rouder (2016) approach focuses on pairwise Bayes factor comparisons between models with the MIC either constrained to be positive, negative or zero. We use the categorical distribution to represent whether the MIC is positive, negative or zero. On the surface, this amounts to only a trivial difference as the Bayes factor can easily be calculated from the categorical priors and *vice versa*. The advantage of our approach is that the categorical distributions afford a hierarchical representation of the MIC category. This allows us to directly examine both the posterior probability that the MIC is a certain category at the group level and at the subject level. Posterior inferences regarding different MIC categories at the individual level possible in principle with Thiele & Rouder (2016) model in which each individual's MIC category is independently sampled from a normal prior distribution. One potential challenge for their approach is that differences across subjects are treated as ratio scale rather than categorical, hence a clear subset of participants with positive MIC and another subset with negative MIC would be treated as uncertain evidence for an average zero MIC.

Ultimately, whether using the Thiele & Rouder (2016) approach or the one we have proposed, we hope that hierarchical Bayesian analyses will allow many more researchers to apply SFT.

---

presentations of their early results at the annual meeting of the Psychonomics Society (Houpt & Fifić, 2013; Thiele, 2014).

6. Acknowledgements

This work was supported by AFOSR grant FA9550-13-1-0087 and NSF grant 1156681.

7. References

Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Sage.

Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrika*, *68*, 7–26.

Dzhafarov, E. N., & Kujala, J. V. (2010). The joint distribution criterion and the distance tests for selective probabilistic causality. *Frontiers in Psychology*, *1*.

Dzhafarov, E. N., & Kujala, J. V. (2014). On selective influences, marginal selectivity, and bell/chsh inequalities. *Topics in Cognitive Science*, *6*, 121–128. URL: `http://dx.doi.org/10.1111/tops.12060`. doi:`10.1111/tops.12060`.

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

Eidels, A., Townsend, J. T., Hughes, H. C., & Perry, L. A. (2015). Evaluating perceptual integration: uniting response-time-and accuracy-based methodologies. *Attention, Perception, & Psychophysics*, *77*, 659–680.

Fifić, M. (2016). Simple factorial tweezers for detecting delicate serial and parallel processes. In J. W. Houpt, & L. M. Blaha (Eds.), *Mathematical Models of Perception and Cognition: A Festschrift for James T. Townsend* (pp. 77–106). New York: Psychology Press volume 1.

Harris, J. (2008). Mindmodeling@home: a large-scale computational cognitive modeling infrastructure. In *Proceedings of The Sixth Annual Conference on Systems Engineering Research 2008* (pp. 246–252). Los Angeles, CA, USA.

Heathcote, A., Brown, S., Wagenmakers, E. J., & Eidels, A. (2010). Distribution-free tests of stochastic dominance for small samples. *Journal of Mathematical Psychology*, *54*, 454–463.

Houpt, J., Townsend, J., & Donkin, C. (2014). A new perspective on visual word processing efficiency. *Acta Psychologica*, *145*, 118–127.

Houpt, J. W., Blaha, L. M., McIntire, J. P., Havig, P. R., & Townsend, J. T. (2013). Systems Factorial Technology with R. *Behavior Research Methods*, *46*, 307–330.

Houpt, J. W., & Burns, D. M. (2017). Statistical analyses for systems factorial technology. In D. Little, N. Altieri, M. Fifić, & C.-T. Yang (Eds.), *Systems Factorial Technology: A Theory Driven Methodology for the Identification of Perceptual and Cognitive Mechanisms*. Cambridge, MA: Academic Press.

Houpt, J. W., & Fifić, M. (2013). A hierarchical approach to distinguishing serial and parallel processing. In *Annual Meeting of the Psychonomic Society*. Toronto, ON.

Houpt, J. W., Heathcote, A., & Eidels, A. (2017). Bayesian analyses of cognitive architecture. *Psychological Methods*, .

Houpt, J. W., MacEachern, S. N., Peruggia, M., Townsend, J. T., & Van Zandt, T. (2016). Semiparametric bayesian approaches to systems factorial technology. *Journal of Mathematical Psychology*, *75*, 68 – 85. doi:http://dx.doi.org/10.1016/j.jmp.2016.02.008.

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., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*, 341–355.

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*, 53–72.

Kirk, R. E. (2012). *Experimental Design: Procedures for the Behavioral Sciences*. (4th ed.). Thousand Oaks, CA: Sage.

Kruschke, J. K. (2010). *Doing Bayesian Data Analysis: A Tutorial with R and BUGS*. Amsterdam: Academic Press.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 520–538.

Ohio Supercomputer Center (1987). Ohio supercomputer center. http://osc.edu/ark:/19495/f5s1ph73.

Ohio Supercomputer Center (2012). Oakley supercomputer. http://osc.edu/ark:/19495/hpc0cvqn.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. URL: http://www.R-project.org ISBN 3-900051-07-0.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374.

Schreiber, B. T., Stock, W. A., & Bennett, W., Jr. (2006). *Distributed Mission Operations Within-Simulator Training Effectiveness Baseline Study: Metric Development and Objectively Quantifying the Degree of Learning*. Technical Report Air Force Research Laboratory Mesa, AZ. URL: `http://www.dtic.mil/dtic/tr/fulltext/u2/a461867.pdf`.

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

Stan Development Team (2014). Stan: A C++ library for probability and sampling, version 2.5.0. URL: `http://mc-stan.org/`.

Stan Development Team (2015). *Stan Modeling Language Users Guide and Reference Manual, Version 2.9.0*. URL: `http://mc-stan.org/`.

Thiele, J. (2014). Bayesian modeling in systems factorial technology. In *Annual Meeting of the Psychonomic Society*. Long Beach, CA.

Thiele, J. E., & Rouder, J. N. (2016). Bayesian analysis for systems factorial technology. URL: `http://pcl.missouri.edu/sites/default/files/p_8.pdf`.

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). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.

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

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*, *39*, 321–360.

Townsend, J. T., & Thomas, R. D. (1994). Stochastic dependencies in parallel and serial models: Effects on systems factorial interactions. *Journal of Mathematical Psychology*, *38*, 1–24.

Van Zandt, T. (2002). Analysis of response time distributions. In J. T. Wixted, & H. Pashler (Eds.), *Stevens' Handbook of Experimental Psychology* (pp. 461–516). New York: Wiley Press volume 4. (3rd ed.).

Yang, C.-T., Chang, T.-Y., & Wu, C.-J. (2012). Relative change probability affects the decision process of detecting multiple feature changes. *Journal of Experimental Psychology: Human Perception and Performance*, . Advance online publication.

Yang, C.-T., Hsu, Y.-F., Huang, H.-Y., & Yeh, Y.-Y. (2011). Relative salience affects the process of detecting changes in orientation and luminance. *Acta Psychologica*, *138*, 377–389.

Yang, C.-T., Little, D. R., & Hsu, C.-C. (2014a). The influence of cueing on attentional focus in perceptual decision making. *Attention, Perception, & Psychophysics*, *76*, 2256–2275.

Yang, H., Fifić, M., & Townsend, J. T. (2014b). Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes. *Journal of Mathematical Psychology*, *58*, 21 – 32. URL: `http://www.sciencedirect.com/science/article/pii/S0022249613001223`. doi:`http://dx.doi.org/10.1016/j.jmp.2013.12.001`.

Zhang, R., & Dzhafarov, E. N. (2015). Noncontextuality with marginal selectivity in reconstructing mental architectures. *Frontiers in Psychology*, *6*.