Applying Systems Factorial Technology to Accumulators with Varying Thresholds

13

Bradley Harding, Vincent LeBlanc, Marc-André Goulet, Denis Cousineau
University of Ottawa, Ottawa, ON, Canada

Before the inception of Systems Factorial Technology (SFT; Townsend & Nozawa, 1995), methods that inferred the information processing architecture present in a cognitive system were based on changes in mean response times (RT), with notable examples being Sternberg's additive method (1969, 1998) and Donders's subtractive method (1969). However, while these methods are monumental additions to experimental psychology, they offered little discrimination power, as various architectures can make similar predictions (also known as mimicking; Snodgrass & Townsend, 1980; Townsend, 1972, 1990). SFT aims to eliminate this uncertainty by using a nonparametric approach that focuses on the entire distribution of RT rather than only the mean RT.

SFT is based on the core assumption that a process is composed of subprocesses that can be *selectively influenced*, i.e., that a subprocess can be significantly impacted without affecting any other subprocesses present. Thus, the most frequent SFT experimental design involves a *Double Factorial Paradigm* where two levels of salience (High – H, or Low – L) are possible for two subprocesses (1 and 2). A complete SFT design requires six conditions: four where both subprocesses are receiving information together at various levels of salience (H_1H_2, H_1L_2, L_1H_2, and L_1L_2), and two conditions where only one subprocess is operating (for example, H_1 and H_2). For a discussion on how to apply this experimental design, consult Houpt, Blaha, McIntire, Havig, and Townsend (2014, pp. 320–326) or the tutorial chapter of this book (Altieri, Fifić, Little, & Yang, 2017).

When data from these six experimental conditions is available, SFT provides an analysis "toolbox" that is used to identify the process's underlying organization. This organization can be characterized by its *architecture*, its *stopping rule*, and its *workload capacity*, assuming *stochastic independence* is met. The *Mean Interaction Contrast* (MIC) and *Survival Interaction Contrast* (SIC) obtained from the first four conditions (H_1H_2, H_1L_2, L_1H_2, L_1L_2) identify the architecture and stopping rule that link both subprocesses. Additionally, *workload capacity* gives information on the process's overall processing capability by comparing its throughput when each subprocess works alone (H_1, H_2) versus when they work simultaneously (H_1H_2). The capacity indicators C_{OR} (for self-terminating and coactive processes) and C_{AND} (for exhaustive processes) describe how combining the subprocesses affects the process's overall performance, be it negatively (limited capacity), positively (supercapacity), or neutrally (unlimited capacity; Townsend & Wenger, 2004). Finally, *stochastic* (*in*)*dependence* provides information on the level of correlation between subprocesses, or the extent to

Systems Factorial Technology, DOI: 10.1016/B978-0-12-804315-8.00016-1

which one subprocess affects the other's capabilities. A way to verify subprocesses' independence is to assess *stochastic dominance* between the four conditions' cumulative distribution functions (Heathcote, Brown, Wagenmakers, & Eidels, 2010) and to confirm that the ordering of RT distributions meets the theoretical expectations. More thoroughly, stochastic dominance is met when the H_1H_2 condition's cumulative RT distribution is always above those of the H_1L_2 and L_1H_2 conditions, which in turn must always be above the L_1L_2 condition. This can also be verified by making sure that the survivor functions of all four conditions never intersect. If any experimental condition falls out of order, stochastic independence does not hold, leading to uninterpretable SFT results; the salience factors manipulated in the factorial design must be re-evaluated to ensure that they are sufficiently different. A main strength behind SFT is that it is nonparametric, giving interpretable results regardless of the RT distribution as long as all assumptions are met and respected. For an entry-level tutorial to the SFT methodology and its analyses, see Harding, Goulet, Jolin, Tremblay, Villeneuve, and Durand (2016).

In Fig. 13.1, we summarize the typical MIC, SIC, and capacity results for each of the five types of organization detected by SFT when both subprocesses are stochastically independent, when selective influence is achieved, and, in the case of capacity, when context invariance holds when expected (Townsend & Wenger, 2004). The first column of Fig. 13.1 helps visualizing the architectures and how they differ from one another.

While SFT was successful in diagnosing processes involved in many cognitive paradigms (such as visual search, Fifić, Townsend, & Eidels, 2008; local–global information processing for people affected by autism, Johnson, Blaha, Houpt, & Townsend, 2010, etc.), one cannot expect it to be accurate if its assumptions are not met. As noted by Eidels, Houpt, Altieri, Pei, and Townsend (2011), models may mimic others' results if selective influence is not respected, leading to uninterpretable results. Additionally, SFT diagnoses must fit what is theoretically expected from the tested data. For example, Townsend and Nozawa (1995) have noted that:

> *Architectures other than serial or parallel can, in some cases, mimic the factorial predictions of these. For instance, a bare Wheatstone bridge predicts underadditivity in an exhaustive processing situation.*
>
> ***(Townsend & Nozawa, 1995, p. 330)***

Both cases above had either a violation of SFT's core assumption or were not simulating a processes' organization that is expected nor can be identified by SFT. In what follows, we present cases of models that do not violate SFT's assumption, are expected and diagnosable by SFT, yet lead to possible misdiagnoses.

Accumulator Models and Threshold Variability

Accumulator models of decision making assume that perceptual mechanisms send activations (often referred to as *evidence*) to a decision process. This process ac-

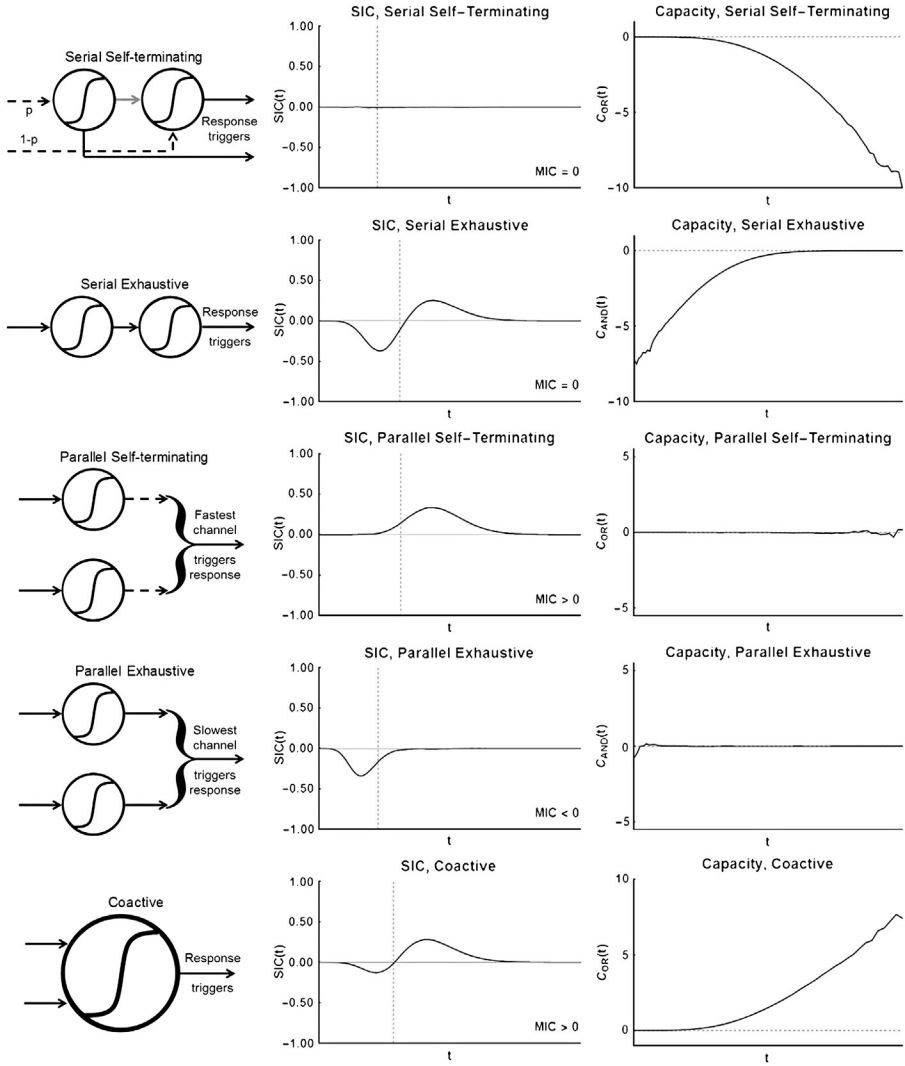

Figure 13.1 Summary of typical SFT results for each of the five information processing architectures assuming stochastic independence between processes (adapted from Townsend & Nozawa, 1995). Column 1 shows schematics of the architectures, column 2 shows the SIC curve along with a typical MIC, and column 3 shows the capacity curves of all five architectures (using the appropriate measure C_{OR} or C_{AND}). In column 1, the circles represent the decision process; the sigmoid curves represent an accumulation-to-threshold implementation of that process; the arrows show the process's input and output (or in the case of serial exhaustive, the shift from one subprocess to another); the dashed arrows denote the fact that only a single subprocess's threshold breach is needed; p is a probability. In columns 2 and 3, the vertical dashed line shows the SIC centerline (see text) and the horizontal dashed line shows the unlimited capacity level.

Table 13.1 LBA threshold parameters values for lower and upper bounds (B and $B + A$ respectively) fitted to four participants (A1, A2, S1, S2) in a lexical decision task. Only the first and last blocks (1 and 25) were estimated using Fig. 4 from Heathcote and Hayes (2012; located on p. 132 in their article)

Participants	Block	A	B	$b = B + A$	Ratio $B : (A + B)$	Proportion $A/(B + A)$
A1	1	1.3	0.1	$0.1 + 1.3$	1: 14.0	93%
	25	0.4	0.1	$0.1 + 0.4$	1: 5.0	80%
A2	1	0.6	0.6	$0.6 + 0.6$	1: 2.0	50%
	25	0.6	0.1	$0.1 + 0.6$	1: 7.0	86%
S1	1	0.6	0.5	$0.5 + 0.6$	1: 2.2	54%
	25	0.6	1.1	$1.1 + 0.6$	1: 1.6	35%
S2	1	2.3	0.6	$0.6 + 2.3$	1: 4.8	80%
	25	0.0	1.4	$1.4 + 0.0$	1: 1.0	0%
Min					1: 1.0	0%
Max					1: 14.0	93%
Geometric mean					1: 3.4	n/a
Mean					1: 4.7	52%

cumulates evidence until a critical amount (a *threshold*) is reached, at which time a decision is made. Accumulator models have a long history in cognitive psychology (e.g., Townsend & Ashby, 1978; Luce, 1986), and their ability to model human RT and response accuracy has been extensively studied (for a recent and complete review of accumulator models, see Forstmann, Ratcliff, & Wagenmakers, 2016; Teodorescu & Usher, 2013). Their flexibility allows the implementation of many factors to model RT distributions; *inter alia*, these models can evolve in discrete or continuous time steps, accumulate discrete or continuous evidence or have either constant or variable thresholds between trials. *Threshold variability* was first introduced in Ratcliff's Drift Diffusion Model (DDM; Ratcliff, 1978) and is a main component in Brown and Heathcote's Linear Ballistic Accumulator (LBA; 2008). It is now commonly used to explain performance variability and the occurrence of fast-errors in RT distributions. This model component will be the focus of our simulations.

In Heathcote and Hayes (2012), the LBA model was fitted to data from a large experiment conducted by Dutilh, Vandekerckhove, Tuerlinckx, and Wagenmakers (2009). In order to grasp the magnitude of their four participants' thresholds variability, we synthesized Heathcote and Hayes (2012) results in Table 13.1, focusing more specifically on the first and last training blocks found in their Fig. 4 on p. 132. In this model, B represents the minimum amount of evidence that must be accumulated to reach the threshold from a starting point that varies uniformly between 0 and A, and b (or $A + B$) represents the maximum amount of evidence that must be accumulated to reach the threshold. We visually estimated B and b, equivalent to the lower and upper threshold bounds.

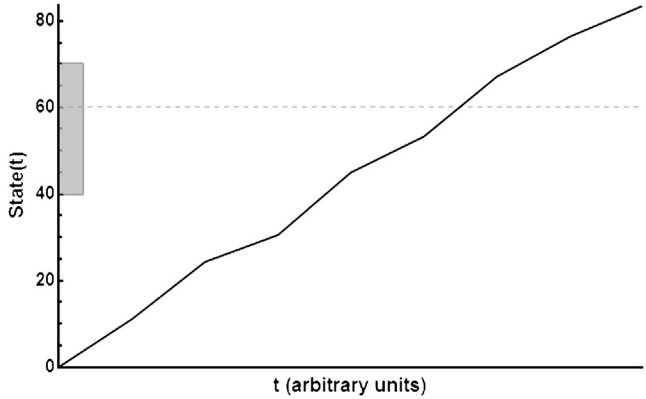

Figure 13.2 Schematic of the DAVT, a discrete accumulator model with varying threshold. In this schematic, the distribution of evidence arrival times is normal with a mean of 4 and standard deviation of 1 (truncated at zero). Threshold variation follows a uniform distribution with bounds at 40 and 70. In this specific trial, the threshold was randomly selected to be 60.

Across trials, the starting point is variable, whereas the threshold (b) is a fixed value. A mathematical equivalent would be to fix the starting point at B and vary the threshold between 0 and A above B. Hence, as few as B evidence may be required on some trials (if the threshold is at 0, its minimal value) to cross the threshold, or as much as $A + B$ (if the threshold is at its maximum value). Therefore, the $B : (A + B)$ ratio is an astute indicator of variability. In this ratio, B represents the minimal threshold value (noted as k_{MIN} in the first model variant presented next) and $A + B$ represents the maximal threshold value (noted as k_{MAX} in the first model variant presented next). A ratio close to 1:1 indicates a negligibly variable threshold whereas a ratio of 1:5, for example, suggests considerable threshold variability (k_{MAX} is five times larger than k_{MIN}). Another way to quantify threshold variability would be to measure the proportion of possible threshold values on the total accumulator's size. This can be done by calculating $A/(A + B)$ where a proportion of 1 indicates that the threshold can vary on the totality of the accumulation's size.

As seen from Table 13.1, the threshold ratio varies greatly between participants and between sessions, going from as low as 1:1 and as high as 1:14, being roughly 1:4 across participants (1:3.4 using the geometric mean or 1:4.7 using the arithmetic mean). Given the magnitude of the threshold variability observed in these empirical results, it is reasonable to explore its impact on SFT diagnostic capabilities.

In this chapter, we investigate SFT's capabilities when analyzing models implementing variable thresholds. For all simulations, we demonstrate that SFT's assumptions are met and the simulated processes' architectures are part of SFT's domain. We do this by simulating the five architectures presented in Fig. 13.1 within two variants of the independent race model class that differ with respect to the locus of the noise (i.e., the stochastic component). The first model variant is a convenient model for demonstration purposes: a stochastically independent discrete accumulator with varying thresholds (hereafter abbreviated to DAVT for readability purposes). DAVT is

Table 13.2 Implementation and parameters of the simulations presented herein. In the top panel we give the parameters of the DAVT. In the bottom panel, we give the implementation equations for all five architectures detected by SFT

List of parameters

Parameters	Description	Value
k_{MIN}	Lower bound of the uniform distribution where threshold values are randomly selected on every trial	(4, 16, 64, 256)
k_{MAX}	Upper bound of the uniform distribution where threshold values are randomly selected on every trial	$(1.25,\ 2,\ 4) \times k_{\text{MIN}}$
β_1	Scale parameter of the Weibull distribution for subprocess 1's evidence arrival times	High = 8 Low = 12
β_2	Scale parameter of the Weibull distribution for subprocess 2's evidence arrival times	High = 10 Low = 20
γ	Shape Parameter of the Weibull distribution for both subprocesses' evidence arrival times	1.8
T_0	Motor time	0

Implementation equations

Architecture	Equation
General equations	$t_i \sim Weibull(\gamma; \beta_i)$ $k \sim Uniform(k_{\text{MiN}}, k_{\text{MAX}})$ $\text{RT}_X = T_0 + \sum_{i=1}^{k} t_i$ where t_i is a random sample of an arrival time distribution and k is randomly selected on every trial for each individual subprocess
Serial Self-Terminating	RT_1 or RT_2 (50% Probability)
Serial Exhaustive	$\text{RT}_1 + \text{RT}_2$
Parallel Self-Terminating	$\min(\text{RT}_1, \text{RT}_2)$
Parallel Exhaustive	$\max(\text{RT}_1, \text{RT}_2)$
Coactive	Identical to the general equations. However, t_i is sampled from two sources of evidence consulted in parallel

schematized in Fig. 13.2 and its implementation is summarized in Table 13.2. The second model is the well-known and documented LBA model (see Table 13.3 for an implementation summary).

Although these models share many similarities, as they are both from the same class of independent race models, they differ in the way they implement their variability. The LBA uses between-trial variability, sampling a different starting point and accumulation rate for each subprocess at each trial. On the other hand, DAVT uses within-trial variability, sampling random arrival times for each individual accumulator.

Table 13.3 Implementation and parameters of the simulations presented herein. In the top panel we give the parameters of the LBA. In the bottom panel, we give the implementation equations for all five architectures detected by SFT

List of parameters

Parameters	Description	Value
A_1	Starting point range of subprocess 1	(1000, 4600, 2300)
A_2	Starting point range of subprocess 2	(1000, 4600, 6900)
b	Threshold	10,000
v_1	Mean accumulation rate of subprocess 1	High = 27 Low = 18
v_2	Mean accumulation rate of subprocess 2	High = 33 Low = 20
s	Standard deviation of the accumulation rate	3
T_0	Motor time	100

Implementation equations

Architecture	Equation
General equations	$v_1 \sim N(v_i; s)$ $a_i \sim Uniform(0; A_i)$ $\mathrm{RT}_X = \frac{b-a_x}{v_x} + T_0$
Serial Self-Terminating	RT_1 or RT_2 (50% Probability)
Serial Exhaustive	$\mathrm{RT}_1 + \mathrm{RT}_2$
Parallel Self-Terminating	$\min(\mathrm{RT}_1, \mathrm{RT}_2)$
Parallel Exhaustive	$\max(\mathrm{RT}_1, \mathrm{RT}_2)$
Coactive	$\mathrm{RT} = \frac{b-a_1-a_2}{v_1+v_2} + T_0$

DAVT is a flexible accumulator and its implementation allows more intuitive controllability over thresholds variation which only necessitates changing the amount of required, summed, arrival times. Hence, it is intuitive when simulating specific threshold ratios (e.g., $k_{\mathrm{MIN}} = 4$, $k_{\mathrm{MAX}} = 8$ means that threshold values can only be 4, 5, 6, 7, or 8 with equal probability). While DAVT is apt at demonstrating proof of concepts, it is not designed to be an applied data-fitting model, i.e., it cannot predict as many facets of human RT as the LBA such as the occurrences of fast errors. Nonetheless, due to its intuitive, discrete accumulation, our simulations will first focus on the DAVT in order to grasp the full effect of threshold variability on SFT diagnoses. Subsequently, we will simulate the five architectures using the LBA to show that DAVT results are generalizable to the stochastically independent race models class.

In the simulations below, we incrementally increased the threshold variability of the race models for all five architectures, and observed whether SFT could still return appropriate results. We used empirically plausible threshold ratios taken from Heathcote and Hayes' (2012) recent model fit of the LBA (1:1.25, 1:2, and 1:4 for the DAVT;

1:1.25 and 1:12.5 for the LBA). Additionally, we verified that the absolute number of evidence required was not responsible for our results. For the DAVT, we did this by varying the lower threshold bound, k_{MIN}, for each threshold ratios ($k_{MIN} = 4$, 16, 64 and 256). For the LBA, we ran one 1:12.5 simulation where both subprocesses had the same value of A (4600), and another with different values of A for each subprocess (2300 and 6900). For all simulations, we assured that the stochastic independence and selective influence requirements of SFT were respected. We expect both model simulations to show similar results as they are from the same class of independent race models. If it is indeed the case, changes in threshold variability, the common source of variability between models, is the component which necessarily affects diagnostic accuracy.

Preliminary Simulations and the Literature on Coactive Architectures

All five architectures in SFT's domain can be implemented within DAVT and LBA. However, the four combinations of serial and parallel architectures with self-terminating and exhaustive stopping rules do not challenge SFT's diagnostic capabilities. While threshold variability increases the processes' decision time variability, diagnostics do not falter as SFT is not sensitive to the amount of variance in the individual processing times.

By contrast, in coactive architectures, threshold variability affects the system as a whole as arrival times from both subprocesses are pooled. For this reason, we will focus solely on threshold variability in the coactive architecture.

Coactive models are much more diverse, and their implementation requires more assumptions compared to serial and parallel models. The original proof describing the SIC curve of a coactive architecture found in Townsend and Nozawa (1995, Theorem 5) is based on a specific implementation of a coactive model, the *Poisson Channel Summation Model* (Schwarz, 1989; Townsend & Ashby, 1983). The same conclusions regarding the shape of the SIC curve were extended to a more general implementation, the *Wiener Coactive Model* (Houpt & Townsend, 2011). However, in both cases, demonstrations were made using sampled accumulation rates (u in Townsend and Nozawa's Theorem, ν_1 and ν_2 in Houpt and Townsend's article). These accumulation rates, within the context of these demonstrations, implied constant thresholds. In Fifić, Little, and Nosofsky (2010), a coactive architecture was implemented within a random walk accumulator model; again, thresholds were set at a fixed value.

These choices were made for simplicity's sake. In the absence of evidence that threshold variability had any impact on coactive diagnoses, there was no point in assuming this parameter could vary. However, the results presented next highlight the importance of this factor.

Results

Simulations were performed with 50,000 trials for each condition (H_1H_2, H_1L_2, L_1H_2, L_1L_2, H_1 and H_2). The survivor functions are shown in the first columns of Figs. 13.3–13.6, the SIC and MIC results (including the 95% confidence interval for MIC) are shown in the second columns, and capacity curves are shown in the third columns. We used the alternative "difference" version of C_{OR} and C_{AND} found in Houpt et al. (2014; for more details, see Houpt & Townsend, 2012). Also presented on all SIC plots are the p-values from a modified Kolmogorov–Smirnov test (Houpt & Townsend, 2010), which detects if the SIC's negative (D^-) and positive (D^+) deviations are significantly different from 0.

The arrival times in DAVT are sampled from a Weibull distribution for implementation purposes. The results presented hereafter are also extendable to other arrival time distributions such as the normal distribution (see Appendix for a demonstration).

DAVT

As seen in the first column of Fig. 13.3, H_1H_2 conditions are the first ordered survivor functions across all k_{MIN} values and the L_1L_2 conditions are always the last ordered. Additional stochastic dominance tests (Heathcote et al., 2010) also indicate that the H_1L_2 and L_1H_2 survivor functions always fall between both extreme conditions ($\Pr(H_1H_2 > H_1L_2) \approx 1$, $\Pr(H_1H_2 > L_1H_2) \approx 1$, $\Pr(H_1L_2 > L_1L_2) \approx 1$, and $\Pr(L_1H_2 > L_1L_2) \approx 1$). Finally, no survivor function crosses any of the other survivor function. When taking the large number of simulations into account, these results indicate that stochastic dominance is present. As seen in the second column of Fig. 13.3, small threshold variability has little to no influence on the interpretation of the SIC nor of the MIC. Our results reproduce what is typically expected of coactive architectures in a SFT analysis. Moreover, the Kolmogorov–Smirnov tests show that the positive and negative portions of all curves are significantly different than 0. Finally, the capacity curves (the third column of Fig. 13.3) indicate supercapacity for all conditions, evidence of the presence of a coactive model. Here, one would make an appropriate diagnosis of the underlying architecture.

In the second set of simulations, where threshold ratio was increased to 1:2, we observe similar results to the ones presented above. In the first column of Fig. 13.4, again, we show that stochastic dominance is met for all four values of k_{MIN}. Moreover, in the second column of Fig. 13.4, the general trend of the coactive SIC curve is still present (an S shape, decreasing below zero, increasing above zero, then returning to zero). However, the entirety of the SIC's amplitude is decreased, with the reduction of the negative portion of the curves being more important. Still, the Kolmogorov–Smirnov tests indicate that all deviations are significantly different than 0.

As threshold variability increases (translated by an increase in the value of the threshold's higher bound, k_{MAX}), so does the possible total duration of the process.

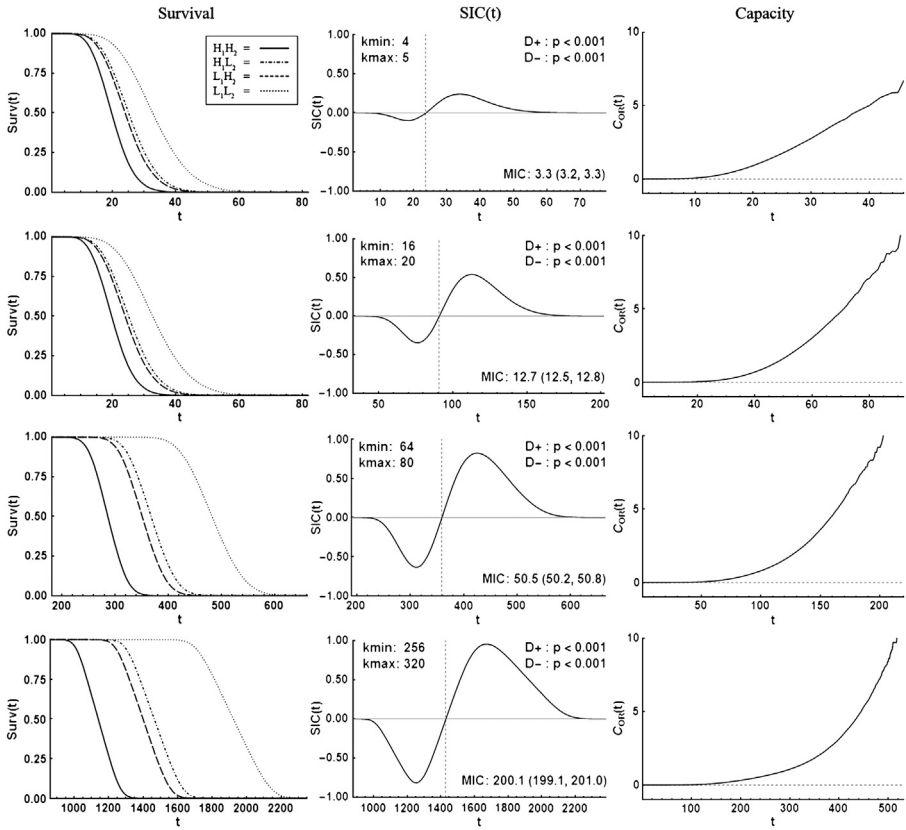

Figure 13.3 Simulation results for simulations using a coactive DAVT with a threshold variability ratio of 1:1.25. All parameters are presented in Table 13.2. The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions. First row, $k_{MIN} = 4$; second row, $k_{MIN} = 16$; third row, $k_{MIN} = 64$; fourth row, $k_{MIN} = 256$.

This results in more variable and platykurtic RT distributions. The survivor functions are therefore less pronounced and drop to 0 much more slowly. As seen, this leads to smaller differences between the survivor functions of H_1H_2, H_1L_2, and L_1H_2, which explains the reduction in SIC amplitude. Additionally, the slower drop of the H_1H_2 survivor functions indicates that the condition benefits less from being organized in a coactive architecture; the overall process's decision time may be extended if the distance between both threshold bounds is at its maximum value. This increase in variability also leads to more overlap between the H_1H_2 and the L_1L_2 survivor function, explaining the later onset of the positive area and the diminishing of the negative area. Finally, the later onset of the positive area combined with the diminishing of the

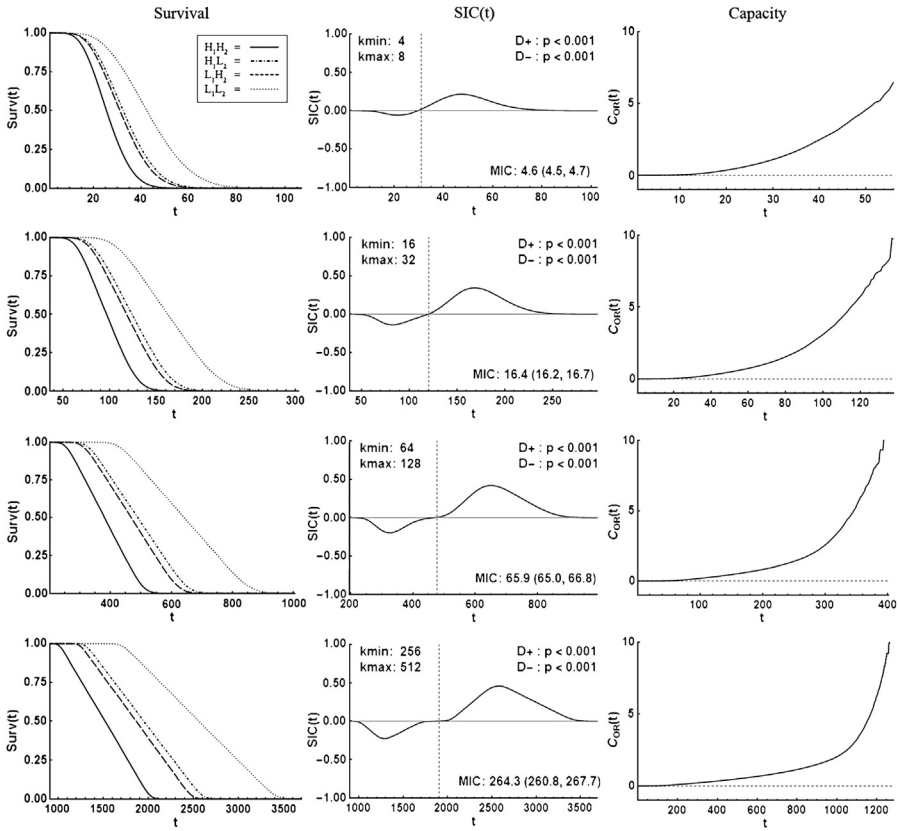

Figure 13.4 Simulation results for simulations using a coactive DAVT with a threshold variability ratio of 1:2. All parameters are presented in Table 13.2. The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions. First row, $k_{MIN} = 4$; second row, $k_{MIN} = 16$; third row, $k_{MIN} = 64$; fourth row, $k_{MIN} = 256$.

negative area creates an atypical "plateau" between the underadditive and the overadditive portions of the curves. This SIC trend depends only on the size of threshold's variability ratio, as it is present across all values of k_{MIN}. Nevertheless, the capacity curves (shown in the third column of Fig. 13.4) still indicate supercapacity. Looking at the SIC and MIC results, the correct coactive diagnosis can be made.

For a threshold variability ratio of 1:4, stochastic dominance remains present (survivor functions are presented in the first column of Fig. 13.5). However, as the increase in ratio increases the amount of possible RT scores, the slopes of all survival functions are further attenuated from both previous simulations. As shown in the second column of Fig. 13.5, the SIC's amplitude is diminished with the negative portion of all SIC

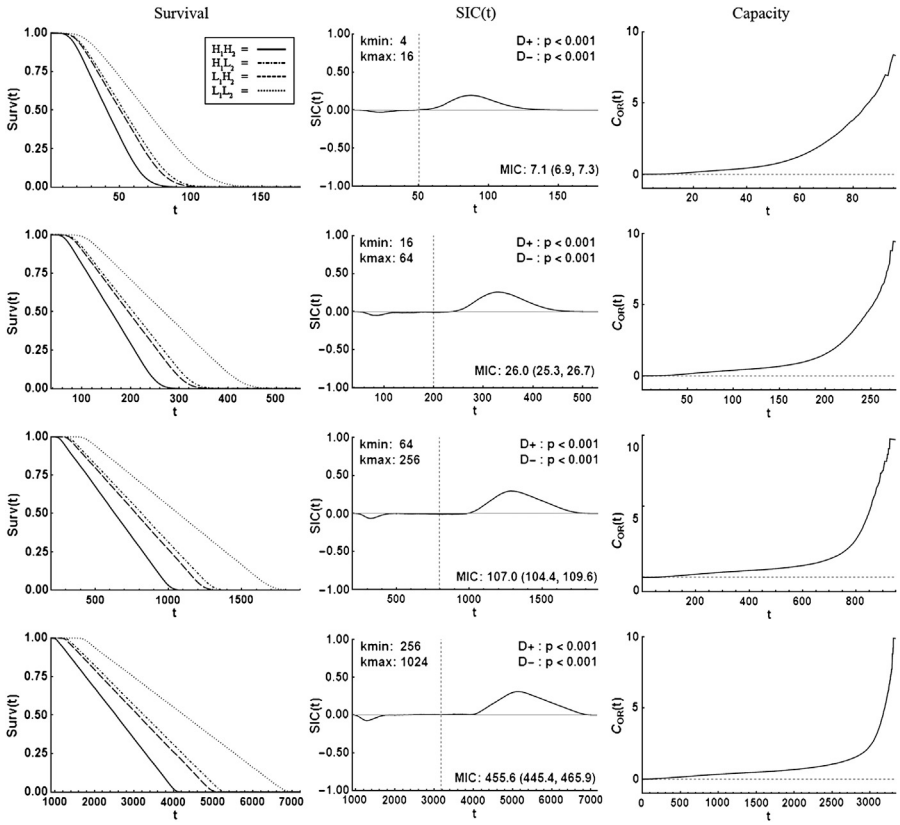

Figure 13.5 Simulation results for simulations using a coactive DAVT with a threshold variability ratio of 1:4. All parameters are presented in Table 13.2. The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions. First row, $k_{MIN} = 4$; second row, $k_{MIN} = 16$; third row, $k_{MIN} = 64$; fourth row, $k_{MIN} = 256$.

curves being almost completely absent from the plot. Results from the Kolmogorov–Smirnov tests still indicate that both negative and positive portions of the SIC curves are significantly different than 0. However, this is only caused by the very large sample size ($n = 50{,}000$, atypical of empirical research) where any small deviation is considered significant. Nonetheless, we believe that if consulted alone, the SIC shape and positive MIC could possibly result in a false parallel self-terminating diagnosis. This possible misdiagnosis is impactful as all of SFT assumptions have been met. Despite these atypical results, as shown in the third column of Fig. 13.5, the capacity curves still indicate supercapacity, which is impossible for a parallel self-terminating process with stochastically independent channels. This highlights the importance of using the entire SFT toolbox to characterize processes.

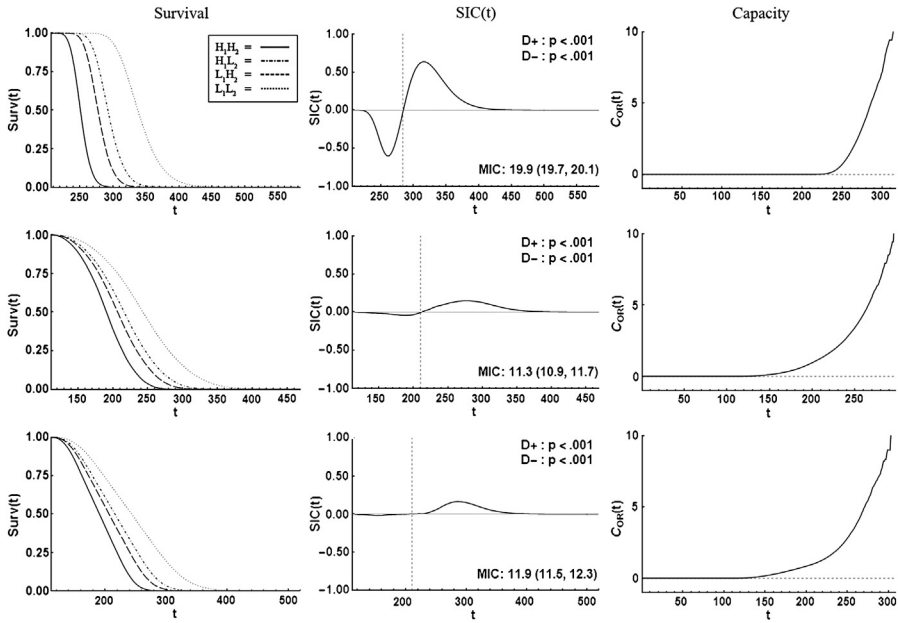

Figure 13.6 Simulation results for simulations using a coactive LBA. All parameters are presented in Table 13.3. The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions. First row, 1:1.25 threshold ratio; second row, 1:12.5 threshold ratio (with even starting points); third row, 1:12.5 threshold ratio (with uneven starting points).

LBA

In order to test whether the presence of threshold variability affects other more complete processing models of RT, we have conducted an SFT analysis on an LBA model (Brown and Heathcote, 2008). In these simulations, shown in Fig. 13.6, we have simulated three LBAs: one with little starting point variability (1:1.25 variability ratio), one with large starting point variability (1:12.5 variability ratio), and one with a large starting point variability and unequal starting point boundaries for each subprocess.

As seen, there is no issue whatsoever with SFTs interpretation of the LBA with little threshold variability. All survivor functions (shown in the first column of Fig. 13.6) follow the expected theoretical order and do not cross at any point. As shown in the second column of Fig. 13.6, the MIC is positive and the SIC follows the expected coactive pattern where the positive and negative portion of the curve are significantly different from 0. Finally, the capacity curve (shown in the third column of Fig. 13.6) correctly indicates supercapacity. From here, we can conclude that a LBA, with little starting point variability, will be correctly diagnosed.

When the starting point variability is large, results are drastically different than the ones presented above and mirror the SIC trend identified using a DAVT. In this sim-

ulation, we can see that the survivor functions are correctly ordered and do not cross at any point. The highlight of this simulation is that the negative area of the SIC curve is greatly diminished. Much like what was found with DAVT, Kolmogorov–Smirnov tests return significant results for both negative and positive portions of the curve, which is due to the very large sample size.

In the third simulation, we observe the same results as with the high threshold variability DAVT simulations. The negative area of the SIC curve is seemingly not present. Once again, Kolmogorov–Smirnov tests show a significant negative and positive portion of the curve. Furthermore, just like it is the case with DAVT, the model is correctly deemed to operate in supercapacity.

These findings highlight the impact of threshold variability on SFT's SIC interpretation of coactive architectures. These atypical SIC findings seem to be generalizable to stochastically independent race models.

SIC Centerline

During the analysis of the simulations presented above, we also developed the *SIC centerline*, a complementary SFT diagnostic tool that helps to visually discriminate between architecture alternatives. The centerline is calculated by pooling all four conditions' RT (assuming that all four conditions have roughly the same amount of trials) and measuring the median of the pooled data. Afterwards, the centerline, representing an RT value, is directly plotted on the SIC curve. Depending on the underlying architecture, the SIC centerline falls in specific areas of the SIC curve. In Fig. 13.1, we see where the SIC centerline should typically be in the SIC plots of all five architectures. In parallel processes, the SIC centerline passes through an area of over/underadditivity, whereas in serial exhaustive architectures and coactive architectures with no threshold variability, the SIC centerline should fall very near (and sometimes directly on) the negative and positive's area of intersection.

When threshold variability increases, as in the 1:2 and 1:4 simulations, the SIC centerline systematically falls in the atypical plateau. Thus, the SIC centerline can also be used to distinguish coactive from parallel self-terminating processes in models using variable thresholds. We observe the same results using both DAVT and LBA, showing the tool's usefulness and generalizability across accumulator models. As shown in Appendix, the SIC centerline is also informative when other distributions of interarrival times are assumed.

Although this measure may not be as reliable as an analysis of capacity, it is worth examining when using SFT. We suggest that if the SIC centerline does not pass near the onset of the over/under-additive portions of the SIC curve nor their point of intersection (if the curve has both negative and positive sections), the possibility of an atypical architecture should be investigated further. There are many ways of calculating the SIC centerline such as the mean of pooled conditions or the median (or mean) of only the H_1H_2 and L_1L_2 conditions. However, we found that the median of pooled RT from all four conditions gave us the most consistent results across simulations and models.

Discussion

Although we only illustrated simulations up to a thresholds ratio of 1:4 using DAVT, and 1:12.5 using the LBA, threshold ratios up to 1:14 (representing a threshold that can cover 93% of the accumulator's total size) have been empirically observed (Heathcote & Hayes, 2012; synthesized in Table 13.1). These results, in tandem, show that when threshold variability increases, regardless of model implementation, the amplitude of the negative area becomes increasingly hard to detect in coactive SIC curves. Thus, the fact that such large ratios are empirically possible (based on model fits) highlights the relevance of this chapter. Additionally, supplementary simulations using DAVT show that, by increasing the threshold ratio to 1:8, the negative portion of the curve visually disappears, and with a threshold ratio of 1:25, the Kolmogorov–Smirnov tests return consistently insignificant results.

This diminishing underadditivity is problematic because the standard coactive SIC signature increasingly resembles the standard parallel self-terminating SIC curve as threshold variability increases. Researchers could possibly make a wrongful diagnosis if their decision is based on MIC and the SIC curve alone. Additionally, as both parallel self-terminating and coactive architectures are somewhat similar in theory (both include two subprocesses working in "parallel" with a decision being made as soon as a sufficient amount of information is accumulated), falsification of one architecture over the other is increasingly difficult. To eliminate this possibility, it is primordial that all possible tools be exhausted.

Moreover, as we wished to provide a proof of concept, our samples sizes were very large to ensure a smooth SIC curve. For a realistic sample size (in the hundreds rather than in the tens of thousands), it may be even more difficult to differentiate models with large threshold variability as the SIC curve will be much noisier. As an example, we simulated a coactive and a parallel self-terminating DAVT, both with a ratio of 1:4 and a $k_{\text{MIN}} = 64$ using 500 simulations rather than 50,000. The results are presented in Fig. 13.7.

As seen in the second column of Fig. 13.7, the model's SICs are easier to confound with one another when the SIC curves are not as smooth. The already attenuated negative portion of the coactive SIC curve, combined with the positive MIC, could be attributed to noise leading to a possible parallel self-terminating misdiagnosis. In fact, a Kolmogorov–Smirnov test indicates that the negative portion of the curve is not significantly different from 0 in both cases, evidence towards a possible erroneous parallel self-terminating diagnosis for the coactive architecture. However, as the capacity curves of both simulated architectures (shown in the third column of Fig. 13.7) are quite different, we can correctly diagnose both architectures assuming that SFTs assumptions are met and respected (which they are in both cases). Furthermore, the SIC centerline falls in a flat portion of the SIC for the coactive architecture rather than in the overadditive area as it is the case in the parallel self-terminating architecture. This observation helps the discrimination between architecture alternatives.

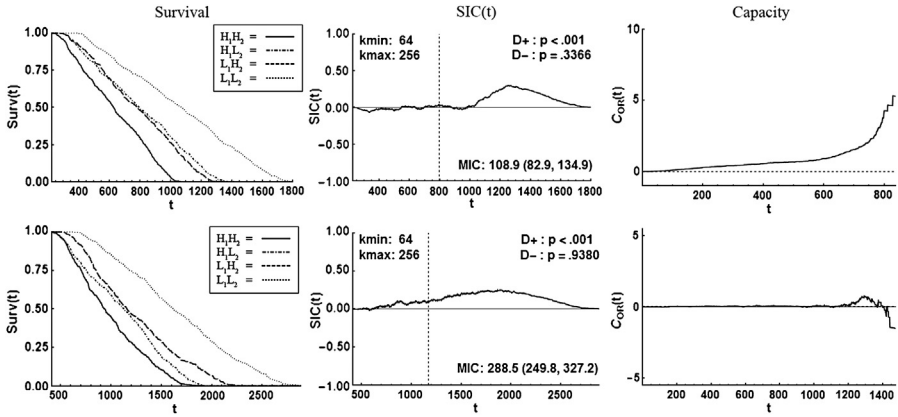

Figure 13.7 DAVT simulation results using $k_{MIN} = 64$, $k_{MAX} = 256$ (1:4 ratio) and $n = 500$ rather than $n = 50{,}000$ (as used throughout this chapter). The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions. First row, a coactive process; second row, a parallel self-terminating process.

Conclusion

As noted by Houpt and Townsend (2011), coactive models are not as broadly defined as the other architectures, limiting the predictions of SIC to the specific cases that have been studied. The present chapter adds to this research and shows that a standard coactive model, with the sole addition of varying thresholds, may make SFT diagnoses on SIC and MIC alone increasingly difficult. These findings were performed with two stochastically independent race model variants of which the only shared feature was threshold variability (or the mathematically equivalent randomly varying starting point) in an environment where selective influence is respected, underlining the impact of our results. This is of importance as very variable thresholds have been used to explain certain participants' RT (Heathcote & Hayes, 2012). Fortunately, we show that the capacity curve eliminates the possibility of user misdiagnosis, a strong argument in favor of using the entire range of the SFT toolbox. We also introduced the SIC centerline, an easy to compute qualitative measure that we developed to assist in diagnosing architectures.

As SFT is fast entering more researchers' repertoires of analyses, it is important to know how varying thresholds, an accepted component to explain human performances (Brown & Heathcote, 2008), affects diagnoses. We believe that with the simulations presented here, one should increase diagnostic prudency and ensure that all tools have been exhausted before making a diagnosis.

Future research should expand on this work and focus on whether SFT can make correct diagnoses of other types of accumulator models with added threshold variabil-

ity, such as the DDM (Ratcliff, 1978), or the Leaky Competing Accumulator model (LCA; Usher & McClelland, 2001).

Acknowledgement

This research project was supported in part by the National Science and Engineering Council of Canada (NSERC).

Appendix: DAVT Simulations Using a Normal Distribution for Evidence Arrival Times

When using normally distributed arrival times of evidence using a normal distribution, it is possible to sample negative arrival times, especially when using small means and reasonable standard deviations. As noted in Donkin, Averell, Brown and Heathcote's (2009), *makedata.r* program, this forces simulated datasets to resample negative arrival times or to truncate them at zero. We avoided this difficulty in our independent race model by using two-parameter Weibull distributions of arrival times. Because it does not require truncation as all the variates are positive or null and because simulation results were identical to the ones with the normal distribution, we opted to keep these distributions. However, to show that the theoretical distribution of the evidence arrival time distributions does not affect SFT results, we replicate in this appendix the simulation with a ratio of 1:4 and a $k_{\mathrm{MIN}} = 64$ using a normal distribution rather than a Weibull as was used throughout DAVT simulations. Implementation is identical to the one presented in Table 13.2 with the sole exception that subprocess arrival times are sampled from the following distributions:

$$\begin{aligned} \mathrm{H}_1 &\sim N(20;5), \qquad \mathrm{H}_2 \sim N(23;5), \\ \mathrm{L}_1 &\sim N(36;5), \qquad \mathrm{L}_2 \sim N(40;5) \end{aligned}$$

where $N(\mu;\sigma)$ is a normal distribution with mean parameter μ and standard deviation parameter σ.

The survivor functions, the SIC (along with the modified Kolmogorov–Smirnov test), the MIC, as well as the capacity curve are presented in Fig. 13.A.

As seen, the results of the simulation are almost indistinguishable from the ones presented using the Weibull distribution. All survivor functions are ordered correctly with no crossing between conditions. Moreover, the negative portion of the curve is attenuated which could lead to a possible parallel self-terminating misdiagnosis. With 50,000 simulated RT, the Kolmogorov–Smirnov tests return significant results meaning that both negative and positive areas of the curve are different from 0. Finally, once again, the capacity curve returns a supercapacity signature. Any evidence interarrival time distribution could be used to replicate the trends stated in the body of this chapter.

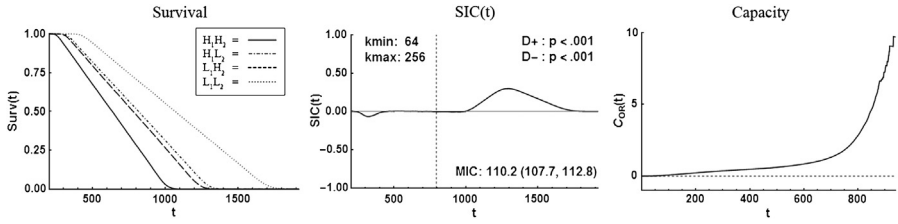

Figure 13.A Results from simulations using a normal distribution rather than a Weibull distribution for evidence arrival times. All DAVT parameters were otherwise identical to the third DAVT simulation (1:4 threshold ratio, $k_{MIN} = 64$) of which the results are shown in Fig. 13.5. The first column shows the survivor functions for all four experimental conditions, the second column shows the SIC curves along with the results of the Kolmogorov–Smirnov tests and the MIC (along with their 95% confidence interval). Finally, in the third column, we show the capacity curves computed using Houpt et al. (2014) measure of C_{OR} based on the difference between cumulative hazard functions.

References

Altieri, N., Fifić, M., Little, D. R., & Yang, C-T. (2017). A tutorial and historical introduction to systems factorial technology. In D. R. Little, N. Altieri, M. Fifić, & C-T. Yang (Eds.), *Systems factorial technology: a theory driven methodology for the identification of perceptual and cognitive mechanisms*. San Diego, CA: Elsevier Publishing (in this book).

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. http://dx.doi.org/10.1016/j.cogpsych.2007.12.002.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. http://dx.doi.org/10.1016/0001-6918(69)90065-1.

Donkin, C., Averell, L., Brown, S., & Heathcote, A. (2009). Getting more from accuracy and response time data: Methods for fitting the linear ballistic accumulator. *Behavior Research Methods*, *41*(4), 1095–1110. http://dx.doi.org/10.3758/BRM.41.4.1095.

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E.-J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, *16*(6), 1026–1036. http://dx.doi.org/10.3758/16.6.1026.

Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology*, *55*(2), 176–190. http://dx.doi.org/10.1016/j.jmp.2010.11.003.

Fifić, 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*(2), 309–348. http://dx.doi.org/10.1037/a0018526.

Fifić, M., Townsend, J. T., & Eidels, A. (2008). Studying visual search using systems factorial methodology with target–distractor similarity as the factor. *Perception & Psychophysics*, *70*(4), 583–603. http://dx.doi.org/10.3758/PP.70.4.583.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Reviews in Psychology*, *67*, 641–666. http://dx.doi.org/10.1146/annurev-psych-122414-033645.

Harding, B., Goulet, M. A., Jolin, S., Tremblay, C., Villeneuve, S.-P., & Durand, G. (2016). Systems factorial technology explained to humans. *The Quantitative Methods for Psychology*, *12*(1), 39–59. http://dx.doi.org/10.20982/tqmp.12.1.p039.

Heathcote, A., Brown, S., Wagenmakers, E. J., & Eidels, A. (2010). Distribution-free tests of stochastic dominance for small samples. *Journal of Mathematical Psychology*, *54*(5), 454–463. http://dx.doi.org/10.1016/j.jmp.2010.06.005.

Heathcote, A., & Hayes, B. (2012). Diffusion versus linear ballistic accumulation: Different models for response time with different conclusions about psychological mechanisms? *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *66*(2), 125–136. http://dx.doi.org/10.1037/a0028189.

Houpt, J. W., Blaha, L. M., McIntire, J. P., Havig, P. R., & Townsend, J. T. (2014). Systems factorial technology with R. *Behavior Research Methods*, *46*(2), 307–330. http://dx.doi.org/10.3758/s13428-013-0377-3.

Houpt, J. W., & Townsend, J. T. (2010). The statistical properties of the survivor interaction contrast. *Journal of Mathematical Psychology*, *54*, 446–453. http://dx.doi.org/10.1016/j.jmp.2010.06.006.

Houpt, J. W., & Townsend, J. T. (2011). An extension of SIC predictions to the Wiener coactive model. *Journal of Mathematical Psychology*, *55*(3), 267–270. http://dx.doi.org/10.1016/j.jmp.2011.02.002.

Houpt, J. W., & Townsend, J. T. (2012). Statistical measures for workload capacity analysis. *Journal of Mathematical Psychology*, *56*(5), 341–355. http://dx.doi.org/10.1016/j.jmp.2012.05.004.

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. http://dx.doi.org/10.1016/j.jmp.2009.06.006.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

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

Schwarz, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, *46*(5), 498–500. http://dx.doi.org/10.3758/BF03210867.

Snodgrass, J. G., & Townsend, J. T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance*, *6*(2), 330–354. http://dx.doi.org/10.1037/0096-1523.6.2.330.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *57*(4), 421–457.

Sternberg, S. (1998). Inferring mental operations from reaction-time: How we compare objects. In D. Scarborough, & S. Sternberg (Eds.), In *Methods, models, and conceptual issues: Vol. 4*. *An invitation to cognitive science* (2nd ed.) (pp. 365–454). Cambridge, MA: MIT Press.

Teodorescu, A. R., & Usher, M. (2013). Disentangling decision models: from independence to competition. *Psychological Review*, *120*(1), 1–38. http://dx.doi.org/10.1037/a0030776.

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 168–199. http://dx.doi.org/10.1111/j.2044-8317.1972.tb00490.x.

Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*(1), 46–54. http://dx.doi.org/10.1111/j.1467-9280.1990.tb00067.x.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In N. J. Castellan Jr., & F. Restle (Eds.), *Cognitive theory, Vol. 3* (pp. 199–239). Hillsdale, NJ: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York, NY: Cambridge University Press.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementation perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*(4), 321–359. http://dx.doi.org/10.1006/jmps.1995.1033.

Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, *111*(4), 1003–1035. http://dx.doi.org/10.1037/0033-295X.111.4.1003.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review*, *108*, 550–592. http://dx.doi.org/10.1037//0033-295X.108.3.550.