Historical Foundations and a Tutorial Introduction to Systems Factorial Technology

Nicholas Altieri1, Mario Fifić2, Daniel R. Little3, & Yang, Cheng-Ta.4

1. Department of Communication Sciences and Disorders, Idaho State University

2. Psychology Department, Grand Valley State University

3. Melbourne School of Psychological Sciences, The University of Melbourne

4. Department of Psychology, National Cheng Kung University

*“Not only is every sensation attended this by a corresponding change localized in the sense-organ, which demands a certain time, but also, between the stimulation of the organ and consciousness of the perception an interval of time must elapse, corresponding to the transmission of stimulus for some distance along the nerves."*

* Abu Rayhan al-Birnuni (c. 973-1048 AD)

*“Time reveals all things”*

* Erasmus

**Introduction**

Conscious experience encompasses a wide variety of rich phenomena: some of which involve the processing of separate sources of information relegated to one sensory modality, and often times, the integration of auditory, visual, tactile, or even olfactory information across sensory modalities.[[1]](#footnote-2) An age-old question in the cognitive and perceptual sciences therefore relates to how the brain processes and combines segregated streams of inputs and unifies them into a conscious experience*.* Even processes that seem rather mundane, such as visually recognizing a tree or a face, or identifying a spoken word, requires a complex cascade of sensory processes and the association of the various forms of information. (For practical purposes, this chapter defines *recognition* as the conscious categorization of an object, sound, or event)

The great Persian scientist al-Birnuni was perhaps the first to notice the interrelationship between temporal and mental processes and task execution. Nonetheless, with the exception of Donders’ subtraction method and Helmholtz’ assays into muscle neurophysiology (i.e., “nerve and muscle physics”; Helmholtz, 1852) formulated in the 19th century, only since the middle of the 20th century have reaction times (RTs) been systematically examined to make inferences about psychological processes. This chapter will briefly summarize some of the major highlights of these fascinating historical developments before providing a tutorial on one of the more recent but seminal developments of RT methodology known as Systems Factorial Technology (SFT) formulated by Townsend and colleagues in the 1990s. [[2]](#footnote-3)

*Examples of Cognitive Processes in the Psychological Literature*

Intra-modal visual and auditory recognition both require intact sensory systems that can process or detect incoming information. This detection and early-stage sensory accumulation process by itself is necessary for recognition, though it is hardly sufficient. To illustrate this point, consider examples of visual agnosia. Patients with visual or other forms of *agnosia*—which essentially translates to *not knowing*—subsequent to stroke or brain injury generally retain the ability to describe the visual or auditory features of a stimulus. What these patients lack is the ability to combine the features in such a way that allows them to understand what they are seeing or hearing. In prosopagnosia, which is a deficit in holistic or configural facial recognition, people lose the ability to identify a face based on information gleaned from seeing individual features such as the eyes, nose, and mouth (e.g., Bauer, 1986). Recognizing a familiar face requires the simultaneous accumulation of information about several different features; however, this is not enough. The information pertaining to the eyes, nose, mouth, and face shape must be somehow associated across feature dimensions or combined in such a way that allows a decision to be made about what face was perceived.

Beyond the scope of recognizing faces (e.g., Wenger & Townsend, 2001, 2006), the basic logic above applies to identifying letters or numbers in a visual display (e.g., Berryhill,, Kveraga, Webb, & Hughes, 2007), recognizing simple stimulus items such as tones or dots (e.g., Miller, 1982; 1986; Miller & Ulrich, 2003; Townsend & Nozawa, 1995), written words (Townsend & Fifić, 2004; Houpt, Townsend, & Donkin, 2014), and even multimodal speech recognition (Altieri, Pisoni, & Townsend, 2011; Altieri & Townsend, 2011; Altieri & Wenger, 2013). An example of multimodal recognition is audiovisual speech perception, such as the McGurk effect; this occurs when listeners are presented with mismatched auditory and visual signals (such as an auditory sound of /ba/ paired with a lip-movement producing “ga”; refer to McGurk & Macdonald, 1976). Often times, the listener will report hearing a fused percept such as “da” or “tha”, rather than the “ba” or “ga” that was actually present.

Several innovative methodologies have been utilized to empirically distinguish between different viable information processing strategies within individual observers. Importantly, these statistical strategies are applicable to various situations and questions in the psychophysical, language, memory, decision making, and vision sciences. These include, but are not limited to: detection of simple visual stimuli (Miller, 1982), change detection (Yang, 2011; Yang, Chang, & Wu, 2013), face recognition (Wenger & Townsend, 2011), multisensory recognition (e.g., Altieri, Stevenson, Wallace, & Wenger, 2015), and so on. This discussion will be accomplished by dissecting processing strategies that describe cognitive processes at a foundational level. We shall see, however, that despite the basic level of these questions, the processes for measurement and computation are highly complex and have undergone considerable theoretical revision over the past century.

The foundational questions that we speak of encompass both *mental architecture* and *workload capacity*. Mental architecture refers to the information processing strategy utilized to; for example, consciously categorize items in a display. Are items—dots, letters, facial features, etc—processed one at a time in a serial manner? Or are they instead processed at the same time in a parallel manner? A subsidiary issue that we shall explore is the decision strategy: this concerns whether all items in a display must be processed before recognition occurs (so called *exhaustive processing*), or instead, whether one can stop and identify the stimulus before all the display items have been processed (*self-terminating processing*). *First-terminating* processing constitute a special case of self-termination, occurring when processing can finish as soon as the first item in the display is correctly identified or otherwise reaches threshold. There are numerous factors which can work to determine the processing strategy; these include several factors which have formed the central research focus of perception and cognition including individual factors (e.g., in learning, Houpt & Blaha, 2015; cognitive ability, Yu, Chang, & Yang, 2014; personality trait, Chang & Yang, 2014), task specific factors (e.g., response biases; Blaha, 2016, this volume), and stimulus-specific factors (e.g., separability and integrality; Griffiths, Blunden & Little, 2016, of this volume; relative saliency, Yang, 2011; Yang et al., 2013).

Next, *workload capacity* deals with whether information processing becomes more or less efficient as the number of items in a display is manipulated. As we shall see, architecture and capacity are logically independent: it is possible, for example, for efficiency or high workload capacity in a serial system, and on the other hand, limited capacity or efficiency in a parallel system. Indeed, quite plausible systems of the latter type have been evoked to explain visual attention processes (Yang, 2016, this volume).

The following section provides details about how the methodology for assessing architecture and capacity has been refined over the past century. We shall see that the methods are complex in the sense that they do not solely rely on obtaining mean accuracy or mean RTs and averaging that data across participants, as is the norm in many experimental paradigms. Instead, the time course of processing is considered at the level of the entire RT distribution typically by contrasting RTs collected for different experimental manipulations. In later sections, we shall demonstrate how the *Double Factorial Paradigm*, or DFP, makes important and strong assumptions all while relying on RT distributions to infer internal information processing strategies.

**Historical Background**

In spite of al-Birnuni’s millennia old idea that temporal processes form an important barometer of cognitive and sensory processes, laboratory work using RTs to infer mental or neurophysiological processes was only commenced in the 19th century. Helmholtz reported physiological studies in the middle of the 19th century in which an electrical shock was administered to the skin, and participants were required to respond by moving their hand as soon as they perceived the shock (Helmholtz, 1850). Importantly, these ideas foreshadowed later developments that subdivided RTs into constituent components including stimulus encoding time, decision time, response selection time, and motor execution time (e.g., Luce, 1986).

Other experiments using RT methodology were carried out by Donders (1969), who devised what became known as the “subtraction method”. The subtraction method is essentially a way to measure processes that occur in a serial fashion. For example, suppose we obtain mean response times from an experiment that requires participants to categorize an object (task A) and respond by choosing between one of two category options (task B), and then we obtain RTs when participants are just asked to respond with either category to some simple stimulation (task B alone). By subtracting the time it takes to complete task B from the total amount of time it takes to complete A and B, we can obtain the estimated time it takes to complete task A alone, the mental time taken for categorization. The assumption underlying the subtraction method is that completion times are strictly additive; however, this is not always true as tasks can interact with one another. In other words, the assumption of strict seriality of certain mental processes does not always hold and must be assessed empirically.

Wilhelm Wundt, the 19th century father of modern psychology, also made forays into the temporal processing domain. According to Wundt’s psychological approach of introspection, complex psychological processes can be reduced to simpler components. Accordingly, and similar to Donders, Wundt’s approach makes the assumption that RTs to complex stimuli should be slower than (i.e., the sum of) simpler stimuli (cf. Robinson, 2001). The goal of more recent statistical methodology in the psychological sciences has been developed to allow us to test the processing assumptions underlying mental processing derived from these forbearers.

One century later, Saul Sternberg’s (1966; 1969) *additive factors method* was developed for the purpose of assessing whether short-term memory search was in fact serial, or alternatively, occurred in parallel; that is, do all stored items from a memory set become activated for recognition simultaneously? In Sternberg’s classic paradigm, participants are given a list of digits to memorize and then shown a “probe” digit after a brief study period. The task for the participant is to answer as quickly and as accurately as possible, as to whether the probe digit was contained in the list of digits. Sternberg’s paradigm included one crucial manipulation: testing what happens to mean RTs when the number of items in the list (i.e., the set size) increases. Hypothetically, as the list of items in short-term memory increases, the time it takes to determine whether the probe is contained in the list should also increase. In a significant development in RT research, Sternberg (1969) found evidence that mean RT does increase as the number of items stored in memory increases and that this increase occurred at the same rate (across set sizes) regardless of whether the probe was presented in the memory set or not. The former result was taken by Sternberg to imply that search occurred in a serial fashion; the latter result was taken to imply that the search did not terminate as soon as the probe was located in the list (i.e., which would result in a decreased rate of increase for target-present trials compared to target-absent trials) but scanned all of the items exhaustively. Together these findings were considered indicative of a serial exhaustive search mechanism

Soon after the development of Sternberg’s (1969) additive factors method, Townsend and colleagues further refined statistically-motivated RT methodology (SFT) to improve the identification of mental architecture. One key limitation of several approaches that use mean RTs, such as the additive factors method, is the problem of “model mimicry”. Model mimicry refers to the myriad of cases in which parallel and serial models can produce identical mean RT signatures. For example, a parallel model with capacity limitations can yield a mean RT slope that increases linearly with set size such as that described by Sternberg (see e.g., Townsend, 1972; Townsend & Ashby, 1983; Wenger & Townsend, 2000). Conversely, a serial model with very efficient processing (i.e., super-capacity) can yield the flat RT set-size function that was commonly believed to be associated with parallel architecture. To further complicate issues, parallel models and coactive models, which pool resources into one common accumulator, can also predict identical mean RT signatures.[[3]](#footnote-4) We refer the interested reader to Townsend and Ashby (1983), Townsend (1990a), and Townsend and Wenger (2004b) for further mathematical and theoretical description of these issues.

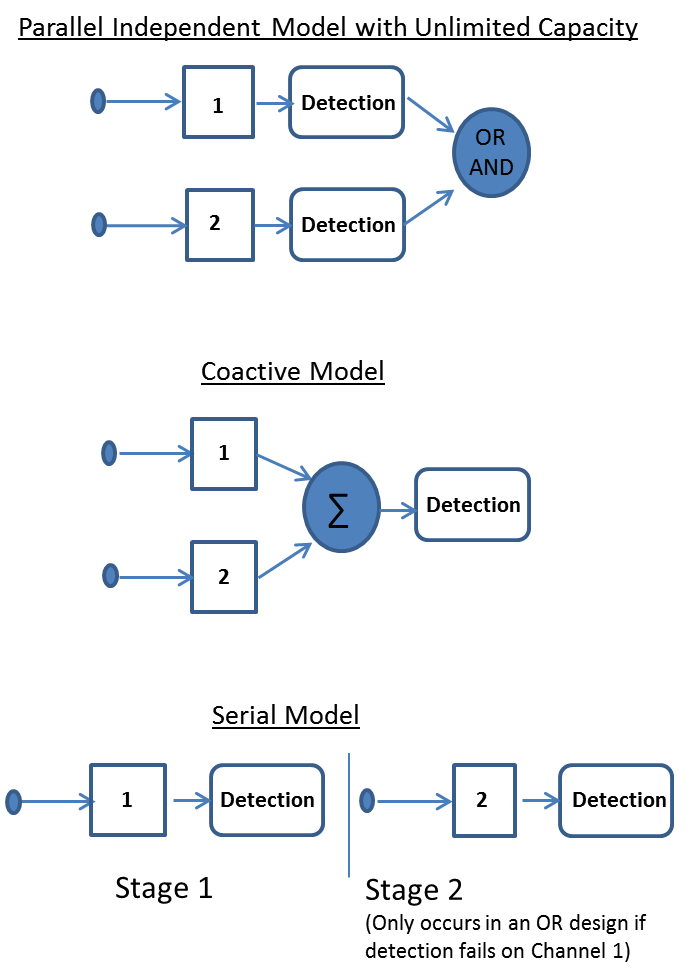
To address this problem of model mimicry, Townsend and Nozawa (1995) developed a more fine-grained theory driven methodology for model testing known as SFT. SFT is a suite or toolbox of methods and statistical tools which improve upon and expand the ability to use RTs to discover important properties of the information processing system. For instance, within SFT, the DFP uses a combined analysis of interaction contrasts across factorial conditions, together with assessments of workload capacity that measure system-level efficiency as a function of workload. These ideas build on the extensive history of Donders, Sternberg, and many others whose work we have not covered here. We direct the reader to the excellent historical reviews by Townsend and Ashby (1983), Luce (1986), Jensen (2006), Schweickert, Fisher and Sung (2012), and Algom, Eidels, Hawkins, Jefferson, and Townsend (2015). For the remainder of the chapter, we focus on explaining precisely what aspects of the processing system we seek to understand, discussing how to construct the DFP, and showing how to use the theoretical SFT tools.

**Properties of information processing systems**

As illustrated by Sternberg’s (1969) considerations, an important perspective in the cognitive sciences is that mental operations must occur in some sequence. While a wide variety of hypothetical cognitive and information processing systems can be devised to account for empirical data, this chapter will focus on three broad classes of models. First, as discussed in relation to Donders’, general serial systems assume that object or feature identification occurs one at a time; importantly, processing on the second item cannot begin until the first item is identified. Another type of architecture is a parallel processing architecture. In a parallel system, items, features, letters, or objects can be processed simultaneously. A third type of architecture is termed a coactive processing architecture (cf. Diederich, 1995; Diederich & Colonius, 2004; Miller, 1982; Schwarz, 1989; Townsend & Nozawa, 1995; Townsend & Wenger, 2004). Coactive systems are similar to parallel systems in many ways. For example, they assume that information processing occurs simultaneously in different channels. Coactive systems differ inasmuch as they assume that the accrued information is pooled into a common processing channel, and hence, the decision is made on the combined information rather than on the individual channels. In each of these “basic” serial and parallel architectures, it is assumed that the processing of each channel proceeds independently. The processing architectures can be made vastly more complex by allowing interactions or cross-talk between the processing channels. Hence, an interactive parallel system might contain facilitatory or inhibitory interactions (Eidels, Houpt, Altieri, Pei & Townsend, 2011; Houpt, Eidels, Altieri, Fifić & Townsend, 2008; Mordkoff & Yantis, 1991). (Similarly, interactions might also occur across channels in serial mechanisms, although this may intuitively appear less plausible. See Townsend & Ashby, 1983).

For serial and parallel systems, one must also consider the issue of the decisional stopping rule. As discussed in the context of Sternberg’s (1969) results, one may intuit that if a system ceases processing as soon as a single item is completed, the RT signature will be different, regardless of the architecture, from cases where all items must be completed before a response is made. First-terminating systems can reach a decision and emit a response time as soon as the first channel accumulates sufficient information. Exhaustive systems, on the other hand, can only emit a response when processing has terminated in each of the channels. Crucially, both serial and parallel systems can be combined with first or self-terminating or exhaustive stopping rule; in other words, architecture is logically independent of decisional rule. Coactive systems differ from parallel and serial models because the exhaustive stopping rule is mandatory. This is due to that fact that coactive models emit a response time when the channel containing all of the combined information reaches its decision threshold. Figure 1 shows a schematic diagram of serial, parallel, and coactive systems in the context of a prototypical detection paradigm with two targets (Townsend & Ashby, 1983; Townsend & Nozawa, 1995; see also Miller, 1982, for an early account of coactive processing using simple auditory and visual stimuli).

SFT also allows one to assess whether the decision is made exhaustively or in a self-terminating fashion. We will focus on the combination of serial and parallel models endowed with an exhaustive or self-terminating stopping rule and on the coactive processing model, for which the question of self-termination is moot. These five models: serial self-terminating, serial exhaustive, parallel self-terminating, parallel exhaustive, and the coactive model form the “Big-5” models of SFT for which theoretical measures are fully developed. More recent work by Eidels et al. (2011) has focused on the characterization of the spectrum of interactive parallel models using the same methods.



*Figure 1*: This is a schematic representation of a parallel independent model (top) with an OR as well as an AND gate; this is similar to the parallel model depicted in Townsend and Nozawa (1995). The coactive model assumes that each channel is pooled into a common accumulator where evidence is accumulated prior to making a decision. Lastly, the figure shows a serial OR model which assumes that processing does not begin on channel 2 until processing completes on channel 1. In an AND design, processing would always begin on channel 2 when processing terminates on channel 1, and detection waits for processing to complete on both channels.

A final aspect of information processing systems concerns how the efficiency of the processing system changes with its processing workload, termed *workload capacity* (or just *capacity,* for short). In general, we consider systems whose capacity can be thought of as limited, unlimited, or even better than unlimited (so-called *supercapacity*). Like the other properties, capacity is logically independent of considerations of architecture, stopping rule, and independence between channels. However, empirically, certain capacity signatures tend to co-occur with certain architectures: Serial systems are usually limited capacity whereas coactive systems are usually supercapacity. Reasonable parallel systems can be limited, unlimited, or supercapacity (see Eidels et al., 2011; Townsend & Wenger, 2004b).

**The Double Factorial Paradigm**

Stated briefly, the DFP involves the factorial manipulation of experimental conditions and the statistical analysis of RT distributions to make inferences about the aspects of mental processes reviewed above. Although other variations are possible, a prototypical double or “redundant-target” detection paradigm presents participants with one, two, or zero stimuli on each trial (see Figure 2). Depending on the task instructions, the participant is required to make a speeded response of one type (e.g., a left button press) when either one or two targets are detected in the display and an alternative speeded response (e.g., a right button press) when no targets are detected in the display. This task is termed an OR-rule task because an affirmative response is made whenever any target is detected on redundant-target trials (e.g., in location 1 *or* location 2), which is when a system can terminate. We contrast this with an AND rule task in which one type of speeded response is made only when two targets are presented (e.g., in location 1 andlocation 2) and the other response is made when one or no targets are presented. This occurs when a system exhaustively analyzes all inputs.

In the DFP the saliency or strength of the targets is manipulated factorially on both the redundant and single-target trials. The goal of this manipulation is to speed up or slow down the RT in each channel. This property has been alternatively referred to stimulus salience (Townsend & Nozawa, 1995) or stimulus discriminability (Fifić, Little & Nosofsky, 2010). In the context of redundant-target trials, for instance, we refer to stimuli in which a *high* detectability target appears in both locations (HH), a high detectability target appears in the first location but a low detectability target appears in the second location (HL), the converse of this situation (LH), and the case in which a low detectability target appears in both locations (LL). In addition, the salience manipulation (L and H) is applied to the conditions in which only one item is presented in either of the location (X) (Figure 2). Hence, the DFP combines both a manipulation of workload by varying the number of possible targets that is useful for assessing information processing capacity, and a factorial manipulation of target detectability that is useful for assessing information processing architecture and stopping rule. We next provide a tutorial introduction to both of these applications starting with the latter assessment of architecture and stopping rule.

C:\Users\littled\Dropbox\20 Years of SFT\Tutorial Chapter\Figure1_Example.tif

*Figure 2*: DFP design showing high- and low-detectability manipulations, along with the redundant-target and single-target trials.

*Assessing Processing Architecture and Decisional stopping Rule*

This section deals with how the DFP can be used to measure architecture. As we shall see, the DFP essentially involves inferring cognitive processes by computing RTs and quantifying potential interactions between experimental factors.

*Selective Influence*

The tools comprising the DFP make an important assumption of *selective influence* (e.g., Dzhafarov, 2003; Schweickert, Fisher, & Sung, 2012; Townsend & Schweickert, 1989). Selective influence implies that there is a strict relationship between the experimental manipulation and the effects of the manipulation on the processes of interests such that an experimental manipulation affects only a single channel (sub-process) within a mental architecture. For example, in a brightness detection task, one has to detect a presence of a stimulus that varies in brightness. The standard finding is that increasing brightness of a stimulus shortens the detection time. The detection time is thought to be composed of several subcomponents (e.g., identification time, decision time, and motor execution time; Luce, 1986). Although it seems natural to believe that the brightness manipulation should only affect the dot detection time, there is no simple way to prove that this is true.

Let’s considered another example in which the task is to detect the presence of any of the two spatially separated dots. In this case, we assume that the cognitive system has to process two channels: one for each dot. Each dot could be presented at either high (H) or low (L) brightness. When testing architecture, ~~Providing that selective influence holds~~, the mean RTs for the high level in a given channel should be faster than mean RTs for the low level in the same channel. The classical method used to test for selective influence in a double-dot task is to determine whether the individual survivor functions, defined as 1-F(t), are properly ordered: *SLL(t)* > *SLH(t),SHL (t)* ≥ *SHH(t)*. The ordering of distributions is a necessary condition for architectural assays within the context of the DFP. Importantly, Townsend (1990b) proved that an ordering of survivor functions (*S(t)*) or cumulative distribution functions (*F(t)*) implies that the means of the distribution are ordered, although an ordering of means does not imply that the survivor functions are ordered.

It is also important to stress that the methodology does not depend on any specific probability distributions or parameters. The relevant data characteristics are predicted by the various classes of architectures (Sternberg, 1969; Schweickert, 1978; Townsend & Ashby, 1983) and are consequently non-parametric. An exception to this is the coactive model predictions, which have been proved for Poisson counting models (although even these predictions happen to be independent of particular parameter values; Townsend & Nozawa, 1995) , Wiener diffusion models (Houpt & Townsend, 2011), and have been shown to hold for discrete random walk models (Fifić, Little & Nosofsky, 2010; Little, 2012; ~~Ratcliff & Smith, 2004~~).

*Mean Interaction Contrast*

One statistic that can be computed using the factorial manipulation is the mean interaction contrast or  (see Sternberg, 1969; Townsend & Nozawa, 1995). In this formula, RTLL is used to denote mean RT of the LL (low-low) detectability target, for example.

The absence of an interaction (i.e., MIC = 0) indicates that the effects of experimental factors are additive—a feature that strongly indicates serial processing regardless of stopping rule. Subsequent theoretical effort led to extensions of MIC tests to parallel and more complex architectures (e.g., Schweickert, 1978; Townsend & Schweickert, 1989; Schweickert & Townsend, 1989; Townsend & Ashby, 1983). While an interaction indicates a lack of evidence for serial architecture, parallel processing cannot be inferred based on a non-zero MIC alone: one notable shortcoming of the MIC mimicry of parallel self-terminating models and coactive models; both predict MIC > 0. Another shortcoming of the MIC is that is that it is a coarse measure only representing one point at each level of detectability (i.e., the mean of the distribution).

*Survivor Interaction Contrast*

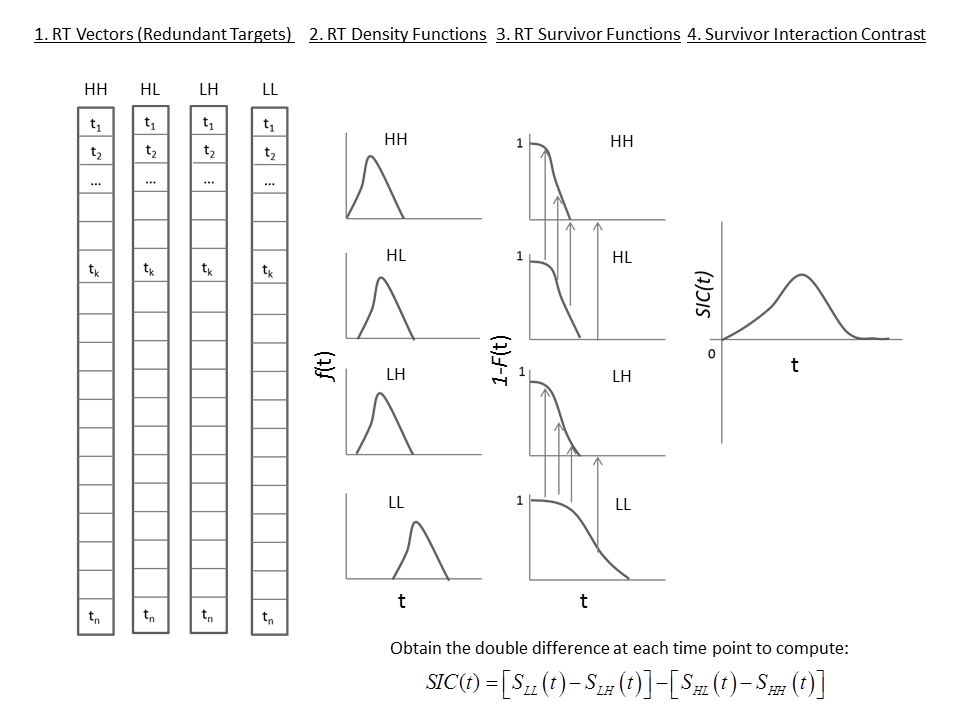
Townsend and Nozawa (1995; also Townsend & Wenger, 2004a for further review) developed a more sensitive contrast that analyzes the functional form of curve of the entire distribution of RTs, namely, the *survivor interaction contrast* (SIC(t)). We may define the SIC mathematically as:



Notice that the SIC(t) uses the same sequence of terms as the MIC, only now survivor functions, S(t), are used rather than the mean RTs. The survivor function *S(t)*, a statistical tool used in *survival analysis* (e.g., Elandt-Johnson & Johnson, 1999), is a function indicating the probability that a process has not completed at time *t*. *S*(*t*) equals one minus the cumulative distribution function, *F*(t), and contains more information that the mean value as it describes the process of interest over the time dimension. Thus, *S(t)* possesses more statistical inferential power than the mean RT (Townsend, 1990b).

Figure 3 shows a diagram of the step-by-step processes involved in computing SIC. First, we obtain vectors of RTs from the LL, LH, HL, and HH experimental conditions for an individual participant. While averaging data across participants is possible, there are both statistical and philosophical issues that can arise when averaging data (e.g., Ashby, Maddox, & Lee, 1994; Estes, 1956; see also Fifić, 2014, for the effect of averaging data on the MIC analysis): One particular drawback is that group averages can obscure important trends arising for an individual participant to such an extent that the average does not resemble any of the individuals. Generally, the number of responses should each contain a large number of trials (i.e., greater than *N* = 100 is a good rule of thumb). Next, from each of these vectors, we empirically calculate the normalized probability density function *f*(t). Third, we obtain the cumulative sum of the *f(t)* values, otherwise known as the empirical cumulative distribution function that is a close estimate of real cumulative distribution function *F(t)*. A simple transformation, namely 1-*F(t)*, yields the survivor function. We refer the reader to Van Zandt (2000) and van Zandt and Townsend (2012) for further details.

The survivor functions should be plotted on the same plot to ensure that they are ordered and that the assumption of selective influence holds. Houpt, Blaha, McIntire, Havig, and Townsend (2014; see also Heathcote, Brown, Wagenmakers, & Eidels, 2010; Houpt & Burns, 2016, this volume) introduce various statistical methods for checking this assumption. If the stochastic ordering of the target stimuli survivor functions holds, then we can compute the interaction at each point of the function to give us the continuous *SIC(t)*. The shape of the function can then be used to diagnose mental architecture.



*Figure 3*: The steps involved in computing the survivor interaction contrast. In the case of the figure, we can surmise parallel processing with an OR decisional stopping rule because the function is over-additive at each point. Of course, the MIC would be greater than 0 as well; however, the SIC(t) gives us more powerful and fine-grained information. The arrows refer to a specific point in time and are provided to aid the visual comparison across the different functions.

*SIC(t) Predictions*

We now turn to the SIC(t) predictions for each of the standard mental architectures and decisional rules (the proofs of the related theorems are presented in Townsend & Nozawa,1995). The SIC(t) predictions for standard parallel, serial, and coactive models with self-terminating and exhaustive stopping rules are shown in Figure 4.

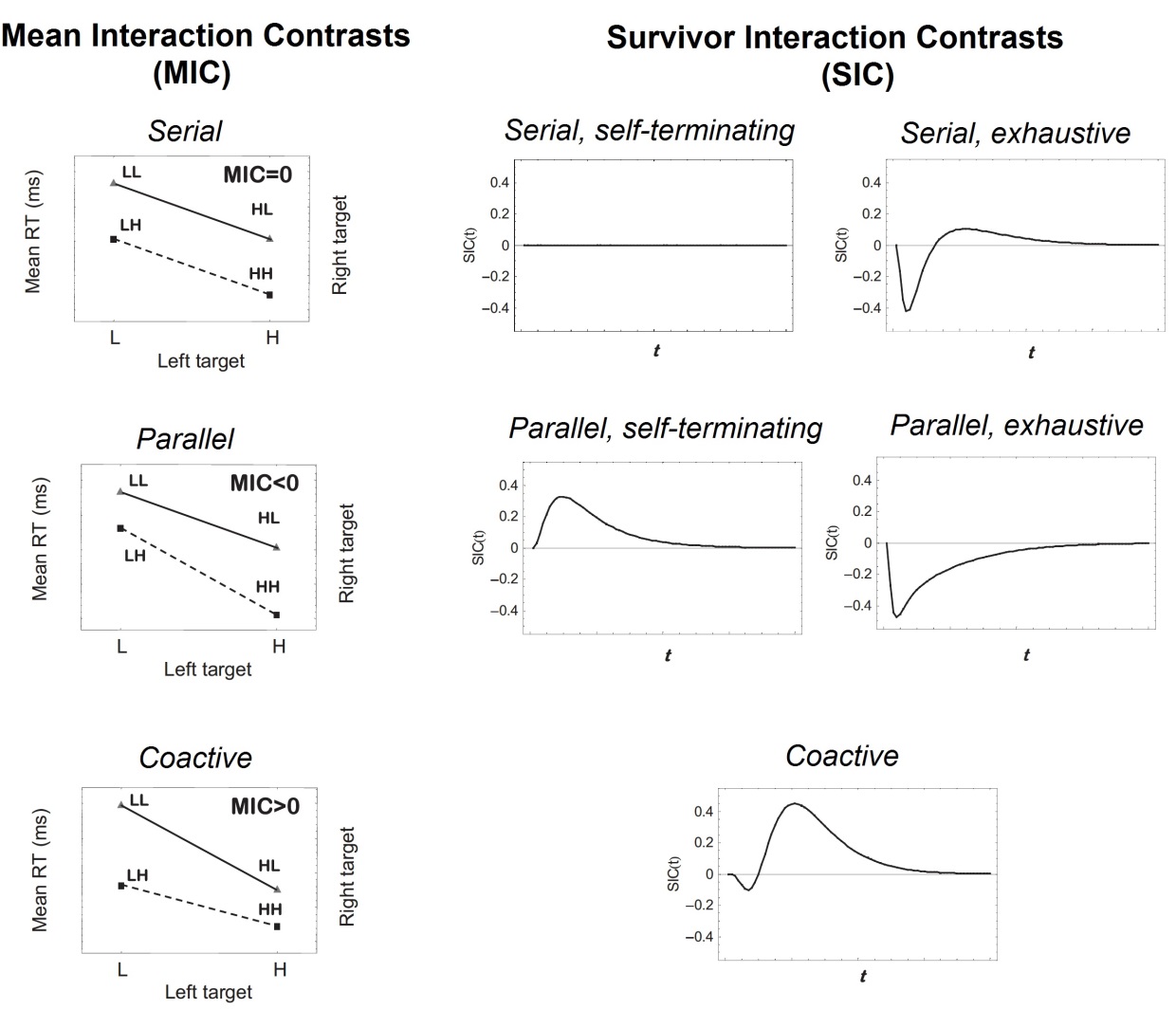
For a parallel self-terminating processing model, the SIC(t) function is entirely positive (i.e., revealing RT over-additivity; see Figure 4) models are appropriate to test when completion of any of the processing channels can correctly decide the response. The intuition for why a parallel self-terminating SIC(t) is positive is because the SLL(t) – SLH(t) term is always larger than the SHL(t) – SHH(t) term in Equation 1. This is because for a parallel self-terminating model, the RT for a redundant target is the minimum time to complete any of the target channels. Hence, the processing time for the LH, HL, and HH stimuli will be much faster than for the LL stimulus because the High detectability component will have a faster RT than a Low detectability component.

A parallel exhaustive model predicts an SIC(t) that is entirely negative (i.e., revealing RT under-additivity, see Figure 4). This exhaustive stopping rule is required in cases where all channels must reach completion before it is certain that a correct response can be made. The intuition for why a parallel exhaustive model predicts a negative SIC(t) is because the SLL(t) – SLH(t) term is always smaller than the SHL(t) – SHH(t) term in Equation 1, for all time points *t*. This is because, in a parallel exhaustive model, the RT for a redundant stimulus is the maximum time necessary to complete any of the target channels. Hence, processing time for the LL, LH, and HL stimuli will be much slower than for the HH stimulus.

A serial self-terminating process predicts a MIC of 0 as does a serial exhaustive process. By contrast, the SIC(t) functions for serial self-terminating and exhaustive processing take on very different shapes. When processing is serial and self-terminating, the SIC(t) is flat and equal to 0 at every point of time (Townsend & Nozawa, 1995). When processing is serial and exhaustive, the SIC(t) is an S-shaped curve with a negative region for early processing times and a positive region for later processing times. The negative and positive regions of the curve are equal to each other in serial exhaustive model, and if we integrate over the curve, the total area is equal to the MIC and must equal zero. Hence, the SIC function delivers strikingly distinct signatures for the important architectures *and* their stopping rules.

Coactive models form a class of parallel models in which the information from each channel is pooled, typically by being added, together into a single channel. While the proofs for SIC(t) functions in coactive models in Townsend and Nozawa (1995) rely on Poisson summation processes, other coactive models such as those based on superimposed diffusion processes, have also been proposed (e.g., Diederich, 1995; Diederich & Colonius, 1991; Miller & Ulrich, 2003; Schwarz, 1989; 1994). Simulations of linear dynamic and Poisson models indicate that the results are general across at least Poisson counter models and Wiener diffusion models (Eidels et al., 2011).

The survivor interaction contrast function for the coactive model is negative at the beginning for the fast RTs and becomes positive at the later or slower RTs. This is similar in shape to the serial exhaustive SIC(t), but note that the initial negative deflection is smaller than the later positive deflection in the coactive model. The MIC for coactive models is positive, similar to self-terminating parallel models. Due to the relation between the survivor function and mean RT (i.e., for RTs,  where  is the mean RT, Townsend, 1997), the integral of the SIC(t) function equals the MIC.

**

*Figure 4*: SIC(t) predictions for independent parallel, serial, and coactive models. The top panels display the predictions of the independent parallel self-terminating and exhaustive models respectively, while the middle panels display the predictions of the serial self-terminating and exhaustive models respectively. The bottom panel displays the coactive predictions. Notice that the predictions of the coactive model appear similar to the serial exhaustive model, the difference being that the predicted size of the negative region in coactive models is smaller than the positive region.

In summary, this section has demonstrated how to use the RTs recorded from redundant targets in a double factorial experiment to construct a factorial contrast test that allows differentiation of different processing architectures and stopping rules. We focused on a detection task, but the double factorial paradigm can be generalized to a wide variety of different domains including recognition memory (Townsend & Fifić, 2004), categorization (see Griffiths et al., 2016; Cheng, Moneer, Christie & Little, 2016, this volume), cued attention (see Yang, 2016, this volume), multimodal processing (Altieri, 2016, this volume), and face recognition (Wenger & Townsend, 2001; 2006). In the next section we turn to the analysis of the remaining component of the DFP; namely, the manipulation of workload across single targets and redundant targets. We describe how information from this manipulation can be used to compute a measure of workload capacity termed the *capacity coefficient.*

*Assessing Workload Capacity*

Assessing a system’s capacity helps answer the question as to whether there is a significant cost, benefit, or no change in processing efficiency as a function of workload. Processing efficiency is essentially determined by comparing processing when multiple processing channels are operating, relative to an unlimited-capacity system whose predictions are derived from single-target trials. Generally, if the processing rate on each channel is unaffected by increasing the number of channels, the system operates at unlimited capacity. If the system slows down on redundant-target trials relative to an unlimited-capacity system, then processing operates at limited capacity. Finally, if the system speeds-up upon the activation of more than one channel, relative to the predictions of an unlimited-capacity system, then the system operates at supercapacity.

As with the architecture tests, capacity predictions and measurements are invariant across specific distributions and parameter values. In addition to often being of interest in its own right, the capacity measure can help arbitrate matters when two distinct architectures and stopping rule combinations yield the same qualitative picture with the SIC(t).

Measuring capacity requires examining the ratio of the integrated hazard functions. To compute this ratio, one could first compute the hazard function as. The hazard function, *h(t)*, indicates the probability that a process will terminate at the next moment in time (*t* + *t*), conditioned on the fact that it has not yet terminated at time *t*. Because, then:



(Luce, 1986, p. 14); hence, the integrated hazard function: . In the classical case of particle arrival at a Geiger counter (Parzen, 1962), the hazard function captures the instantaneous and time invariant likelihood of a particle’s arrival at any point in time given non-arrival before that point.  *H(t)* is a slightly coarser measure of efficiency, or work completed by time *t*, since it integrates h(t); importantly however, it still carries with it inferential statistical advantages over means (Townsend, 1990b). Further, in estimation from data, it has been our experience that it tends to smooth out irregularities and provide more robust estimates.

To empirically calculate the capacity coefficient at each time interval, the integrated hazard functions must be calculated for the conditions in which the participant is presented with redundant-target information and divided by the sum of the integrated hazard functions obtained from the single-target conditions. The subscripts “1” and “2” in the following equation denote the information presented in each of the two processing channels:

 .

Capacity assesses performance in a system when redundant targets are present versus when only a single target is present by using predictions derived from independent parallel models with unlimited capacity (UCIP). In an OR rule, the UCIP model, in addition to assuming unlimited capacity and stochastically independent processing times, also assumes a self-terminating stopping rule. The standard parallel model thus acts as a benchmark for measuring workload capacity and predicts C(t) = 1, for all *t* > 0. Note that alternative benchmark models could also be used (see for instance, Houpt et al., 2014); however, the parallel model is convenient due to its historical precedence in devising other tests of processing (e.g., the role of the race model inequality found in Miller, 1982).

Capacity outcomes can be divided broadly into three categories based on the assessment of its values into: *Unlimited capacity*, C(t) = 1, means that RTs obtained from redundant-target trials equal UCIP predictions; *Supercapacity*, C(t) > 1, occurs when a parallel model, of equal efficiency with a workload of n items to process, will have one or more of its channels speeding up when compared to cases with a workload of n-1 or smaller. In other words, it means that the amount of work done on, say, two units is smaller than the sum of the work done on each unit separately; Finally, *limited capacity*, C(t) < 1, indicates that processing is less efficient compared to baseline UCIP models.

An upper bound on performance for separate activation parallel models, also known as *independent race models* or UCIP modelswas provided by Miller (1982). It stipulates that in such models, it must be the case that. In this equation,  represents the cumulative distribution function (CDF) for the redundant-target trials, and  andare the corresponding distributions for the single-target trials (e.g., when only one visual stimulus is presented). If C(t) > 1 for even a small interval early in processing then the above inequality has to be violated. And conversely, if there are values of *t* where then it must be the case that  for any such values.

Coactive models possess a strong propensity to exhibit supercapacity: This tendency was one of the reasons for the advent of the concept, first primarily in an operational form. Models were classified as “faster than race models” if Miller’s race inequality was violated. Quantitative interpretations began to appear (Colonius & Townsend, 1997; Diederich & Colonius, 1991; Miller, 1991; Schwarz, 1994; Townsend & Nozawa, 1995). The Townsend and Nozawa (1995) study proved that a wide variety of coactive models based on counting processes with arbitrary distributions (including, but not limited to the Poisson distributions) imply not only, but also that such models will at some point violate the race inequality. Subsequently, Houpt and Townsend (2011) demonstrated supercapacity for coactive Wiener processes.

Limited capacity occurs when  (Neufeld, Townsend, & Jette, 2007; Townsend & Nozawa, 1995; Wenger & Townsend, 2000). Limited capacity could easily result from the allocation of a limited capacity source to the various operational channels. A special case of interest occurs when the processing source, measured by the integrated hazard function, is fixed. The system is then said to operate at *fixed capacity* and the  will be the average of the two single-target integrated hazard functions  with *p* lying between 0 and 1 according to the axioms of probability theory. An intuitive example is found if we assume equal distribution parameters with; this implies that the system operates at so called *fixed capacity* when C(t) = ½.

In the same way that the race model inequality places an upper bound on the level of supercapacity achievable by a UCIP model, there is an analogous bound on limited capacity termed the “Grice bound” (Grice, Canham, & Gwynne, 1984). This bound assumes that the fastest of the two channels (or items) is slower that responses to the redundant-target condition; it is expressed as *MAX*{F1(t), F2(t)} ≤ F12(t). (We refer the interested reader to Townsend and Eidels, 2011 for translations of upper and lower bounds into capacity space). When redundant-target processing speed is lower than this bound, we say that capacity is “extremely limited”. Furthermore, if F1(t) = F2(t) = F(t) in a parallel model, then the system operates at “fixed capacity”. Standard serial models also make the same prediction, which highlights the importance of distinguishing the concepts of “capacity” and “architecture.

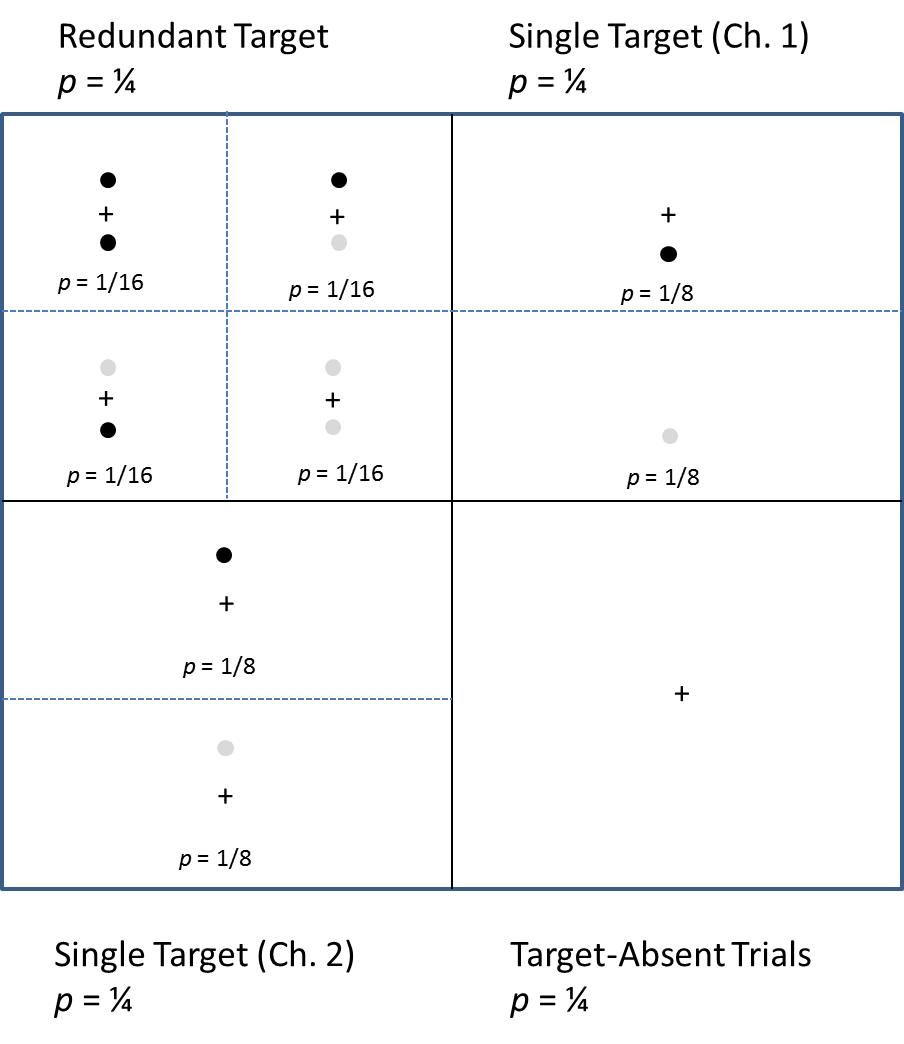
The UCIP baseline model described above clearly provides a benchmark of performance when processing can self-terminate and still respond accurately. This would be true when the task responses were defined by an OR decision rule; consequently, the capacity coefficient defined in Equation is sometimes denoted by. By contrast, in an AND decision task, both target locations must be processed exhaustively to ensure accurate detection of the double target. Termination after processing only a single target would, on some trials, results in inaccurate responding. In this situation, capacity is measured relative to an alternative baseline: a parallel exhaustive model, which expects the double-target processing time to be consistent with the maximum time taken to process either of the targets (i.e., ). Hence, capacity for an AND task is computed as:



where is the integrated reverse hazard function, which is defined, analogously to  as . Defining capacity in this way and inverting the locations of the single and double targets in the numerator and denominator of the function (compare Equations (2) and (3)) has the effect of (a) measuring capacity relative to a UCIP model with an exhaustive stopping rule and (b) maintaining all of the expected relationships from the computation of  regarding limited capacity: , and supercapacity, .

*Probabilistic contingencies*

Finally, another important component of DFP design concerns the avoidance of probabilistic contingencies. In order to prevent probabilistic information from facilitating or alternatively inhibiting redundant-target speed relative to the single-target trials, it must be impossible for the observer to predict the occurrence of a target in channel 1 given the presence of a target in channel 2, and *vice versa* (cf. Mordkoff & Yantis, 1991).Suppose for instance, the probability of a target appearing in channel 2 increases with a target appearing in channel 1; here, redundant-target processing speed will be faster than single-target trials. One straightforward way to avoid probabilistic contingencies within the context of a simple go/no-go visual detection paradigm is to equate the number of redundant-target, single-target, and target-absent trials. Such a scenario is shown in the example OR detection design shown in Figure 5, which requires the observer to say “yes” when a dot appears either above or below a central fixation and withhold their response in the target absent condition.

**

*Figure* 5: Trial types and probabilities for the redundant-target, single-target, and target-absent conditions in a DFP design in which contingencies are avoided. The black dot represents high-salience trials, whereas the lighted gray dots represent low-salience trials.

**Conclusion**

In this chapter, we introduced and reviewed two of the major theoretical contributions arising from Townsend and Nozawa’s (1995) introduction of Systems Factorial Technology. This included an exploration of architecture and capacity measures, accompanied by a brief presentation of the prototypical experimental design. Numerous other extensions exist, and relevant adaptations continue to appear in the psychological literature: For example, a promising line of neuroimaging research has extended the concept of additive factors from the realm of RTs to neural-biological markers (Stevenson & James, 2009), as well as effects of cross-channel interactions on architecture and capacity (see Eidels et al., 2011; see also Wenger & Townsend, 2004b for effects on capacity).

Additionally, numerous theoretical developments, many of which were highlighted in this chapter deal with extensions of theoretical issues (Townsend & Liu, 2016; Algom & Eidels, 2016, this volume), the capacity function (Cheng et al., 2016), the double factorial paradigm for three or more channels (Fifić, 2016a; Yang, Fifić & Townsend, 2014), statistical measures applicable to systems factorial technology and the survivor interaction contrast (e.g., Houpt & Townsend, 2010; Houpt & Burns, 2016, this volume), applications in cognitive brain science, and quantitative relationships to other statistical measures will be discussed. The remainder of the book focuses on many of the important extensions to methodology and theory arising from this work along with the application of SFT.

First, the capacity measure described in this introductory chapter only is applicable to correct decisions and only to the case where the single targets are presented in isolation. Altieri’s (2016) chapter on the capacity assessment function describes an extension of capacity to also account for accuracy and for RTs generated from error responses. In this volume, Cheng, Moneer, Christie, and Little’s (2016) describe an extension of capacity to account for the case where distractors are presented along with the single targets. Blaha’s (2016) chapter describes how specific task demands might influence the observed capacity measure.

To facilitate the methodological growth of the SFT applications a general approach for understanding methodological SFT manipulations is proposed in Fifić’s (2016) chapter in this volume. The chapter provides an overview of different SFT applications across various psychological modalities and cognitive tasks, and shows how to integrate the seemingly different type of factorial process RT stretching manipulations into one framework.

In terms of new statistical developments, Houpt and Burn’s (2016) chapter in this book proposes a number of statistical tests that can be utilized to determine whether the capacity coefficient, C(t), is significantly greater or less than UCIP predictions, or whether two capacity coefficients differ significantly from each other.

Finally, the remaining chapters shall review specific applications of SFT beyond simple detection tasks. Some of these include applications in multisensory speech perception (Altieri, 2016b), categorization (Griffiths et al., 2016), auditory processing (He, Lentz, & Townsend’s 2016), selective responding (Algom & Eidels, 2016), attention (Yang, 2016), and configurality (Wenger, Ingvalson, & Rhoten, 2016). Appropriately, links between SFT and related paradigms such as the simultaneous-sequential paradigm (Howe & Ferguson, 2016), multidimensional scaling (Eidels & Howard, 2016), parametric modeling (Hardin et al., 2016; Griffiths et al., 2016), and clinical neuroscience (Taylor, Theberge, Williamson, Densmore, & Neufeld, 2016) will also be addressed in this volume.

**References**

Algom, D. & Eidels, A. (2016). Relation of SFT architectures to selective responding: Associations and dissociations. 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*. Elsevier.

Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of Response Times: Identification of Cognitive Mechanisms through Mathematical Modeling. *The Oxford Handbook of Computational and Mathematical Psychology*, 63.

Altieri, N. (2016a). Development and applications of the capacity-function that measures accuracy. 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*. Elsevier.

Altieri, N. (2016b). An introduction to the double factorial paradigm: How it can answer questions for multisensory researchers. 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*. Elsevier.

Altieri, N, Pisoni, D.B., & Townsend, J.T. (2011a). Behavioral, clinical, and neurobiological constraints on theories of audiovisual speech integration: A review and suggestions for new directions. *Seeing and Perceiving*, *24*, 513-539.

Altieri, N., & Stevenson, R.A., Wallace, M.T., & Wenger, M.J. (2015). Learning to associate auditory and visual stimuli: Capacity and neural measures of efficiency. Brain Topography, 28(3), 479-493.

Altieri, N. & Townsend, J.T. (2011). An assessment of behavioral dynamic information processing measures in audiovisual speech perception*. Frontiers in Psychology, 2*(238), 1-15*.*

Altieri, N., & Wenger, M. (2013). Neural dynamics of audiovisual integration efficiency under variable listening conditions: an individual participant analysis. *Frontiers in Psychology*, *4*(615), 1-15.

Ashby, F.G., Maddox, W.T., & Lee, W.W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science, 5*(3), 144-151.

Bauer, R.M. (1986). The cognitive psychophysiology of prosopagnosia. In H. Ellis, M. Jeeves, F. Newcombe & Andrew W. Young (eds.), Aspects of Face Processing**.** Martinus Nijhoff: 253-267.

Berryhill, M., Kveraga, K., Webb, L., & Hughes, H. C. (2007). Multimodal access to verbal name codes. *Perception & Psychophysics,* 69(4), 628-640.

Blaha, L. (2016). An examination of task demands on the elicited processing capacity. 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*. Elsevier.

Chang, T.-Y., & Yang, C.-T. (2014). Individual Differences in Zhong-Yong tendency and Processing Capacity. *Frontiers in Psychology, 5*.

Cheng, X. J., Moneer, S., Christie, N. & Little, D. R. (2016). An information processing perspective to understanding conflicting information. 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*. Elsevier.

Colonius, H. & Townsend, J.T. (1997). Activation-state representation of models for the redundant-signals-effect. In A. A. J. Marley (Ed.), *Choice, Decision and Measurement*, volume in honor of R. Duncan Luce, Mahwah, NJ: Erlbaum Associates.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology, 39*(2), 197-215.

Diederich, A., & Colonius, H. (1991). A further test of the superposition model for redundant signals effect in bimodal detection. *Perception & Psychophysics, 50*, 83-86.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement of reaction time: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics, 66*(8), 1388-1404

Donders, F.C. Over de snelheid van psychische processes. Onderzoekingen gedann in het Physiologish Laboratorium der Utrechtsche Hoogeschool, 1868-1869, Tweede Reeks, II, 92-120. In W.G. Koster (Ed. And translator), *Attention and Performance II*. Amsterdam: North-Holland, 1969. (Reprinted from *Acta Psychologica*, 1969, *30*, 412-431).

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

Eidels, A., Houpt, J., Altieri, N. Pei, L., & Townsend, J.T. (2011). Nice guys finish fast and bad guys finish last: A theory of interactive parallel processing. *Journal of Mathematical Psychology, 55*, 176-190.

Eidels, A. & Howard, Z. (2016). Relating Systems Factorial Technology to other measures of human performance. 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*. Elsevier.

Elandt-Johnson, R. & Johnson, N. (1999). Survival Models and Data Analysis. New York: John Wiley & Sons.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological bulletin*, *53*(2), 134 -140.

Fifić, M. (2014) Double jeopardy in inferring cognitive processes. *Frontiers Psychology*. 5:1130.

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, 309-348.

Fifić, M., & Little D. R. (2016). Stretching Mental Processes: An Overview of and Guide for SFT Applications. To appear 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. Elsevier.

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

Grice, G.R., Canham, L. Gwynne, J.W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics, 36*, 565-570.

Griffiths, D. W., Blunden, A. G. & Little, D. R. (2016). Logical-rule based models of categorization: Using Systems Factorial Technology to understand feature and dimensional processing. 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*. Elsevier.

Harding, B., Goulet, M-A., LeBlanc, V., Trembley, C. & Cousineau, D. (2016). Applying Systems Factorial Technology to discrete accumulators with varying thresholds. 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*. Elsevier.

He, Y., Lentz, J. J. & Townsend, J. T. (2016). Are two ears always better than one? The capacity function says no. 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*. Elsevier.

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.

Helmholtz, H. (1850). Vorläufiger Bericht über die Fortpflanzungsgeschwindigkeit der Nerventhätigkeit. Archive of the Berlin Brandenburgische Akademie der Wissenschaften.

Houpt, J. W., & Blaha, L. M. Exploring Individual Differences via Clustering Capacity Coefficient Functions.

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.

Houpt, J. W. & Burns, D. (2016). Statistical analyses for 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*. Elsevier.

Houpt, J. Pei, L. Eidels, A., Altieri, N., Fifić, M. & Townsend, J.T. (2008). *An exploration of parallel-interactive information processing models*. Paper presented at the 41th Annual Meeting of the Society for Mathematical Psychology, Washington D.C.

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. (2011). An extension of SIC predictions to the wiener coactive model. *Journal of mathematical psychology*, *55*(3), 267-270.

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

Howe, P. D. L. & Ferguson, A. (2016). The identity-location binding problem. 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*. Elsevier.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Little, D. R. (2012). Numerical predictions for serial, parallel, and coactive logical rule-based models of categorization response time. *Behavior research methods*, *44*(4), 1148-1156.

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

Marrot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain and Language, 79*, 309-320.

McGurk, H., & MacDonald, J.W., (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247-279.

Miller, J. (1986). Time course of coactivation in bimodal divided attention. *Perception & Psychophysics, 40*(5), 331-343.

Miller, J., & Ulrich, R. (2003). Simple reaction time and statistical facilitation: A parallel grains model. *Cognitive Psychology, 46*, 101-151.

Mordkoff, J.T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 520-538.

Neufeld, R.W.J., Townsend, J. T., & Jetté, J. (2007). Quantitative response time technology for measuring cognitive-processing capacity in clinical studies. In R.W.J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms.* Washington, D.C.: American Psychological Association.

Parzen, E. (1962). On estimation of a probability density function and models. *The annals of mathematical statistics*, 1065-1076.

Robinson, D.K. (2001). Reaction time experiments in Wundt’s institute and beyond. In R.W. Rieber and D.K. Robinson (Eds.), *Wilhelm Wundt in history: The making of scientific psychology* (pp. 161-204). New York: Kluwer/Plenum.

Schweickert, R., Fisher, D. L., & Sung, K. (2012). *Discovering cognitive architecture by selectively influencing mental processes* (Vol. 4). World Scientific.

Schweickert, R., & Townsend, J. T. (1989). A trichotomy method: Interactions of factors prolonging sequential and concurrent mental processes in stochastic PERT networks. *Journal of Mathematical Psychology, 33,* 328-347.

Schwarz, W. (1989). A new model to explain the redundant-signals effect. Perception & Psychophysics, 46, 498–500.

Schwarz, W. (1994). Diffusion, superposition, and the redundant targets effect.Journal of Mathematical Psychology, 38, 504–520.

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

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

Stevenson, R.A., & James, T.W. (2009). Neuronal convergence and inverse effectiveness with audiovisual integration of speech and tools in human superior temporal sulcus: Evidence from BOLD fMRI. *NeuroImage, 44*, 1210-1223.

Taylor, R., Theberge, J., Williamson, P., Densmore, M. & Neufeld, R. W. J. (2016). Using SFT to elucidate the “f” of clinical fMRS (functional Magnetic Resonance, Spectroscopy). 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*. Elsevier.

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.

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

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

Townsend, J.T., & Ashby, F. G. (1983). *The Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.

Townsend, J.T., & Eidels, A. (2011) Workload Capacity Spaces: A Unified Methodology for Response Times. *Psychonomic Bulletin & Review, 18*, 659-681.

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

Townsend, J. T. & Liu, Y. (2016). Selective influence and categorical separability (perceptual separability) in perception and cognition: Similarities and distinctions. 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*. Elsevier.

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*(4), 321-359.

Townsend, J.T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology, 33,* 309-327.

Townsend, J.T., & Wenger, M.J. (2004a). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review, 11*, 391-418.

Townsend, J.T., & Wenger, M.J. (2004b). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review, 111*(4), 1003-1035.

van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review,* *7*(3), 424-465.

Van Zandt, T., and Townsend, J. T. (2012). Designs for and Analyses of Response Time Experiments. In Little, T. K. (ed.), *The Oxford Handbook of Quantitative Methods*. New York: Oxford University Press.

Wenger, M. J., Ingvalson, E. M. & Rhoten, S. E. (2016). The evolution of systems factorial theory and evolving conceptions of configurality. 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*. Elsevier.

Wenger, M.J., & Townsend, J.T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *The Journal of General* *Psychology, 127*(1), 67-99.

Wenger, M. J., & Townsend, J. T. (2001). Faces as gestalt stimuli: Process characteristics. In M. J. Wenger & J. T. Townsend (Eds.). Computational, geometric and process perspectives on facial cognition: Contexts and challenges. Lawrence Erlbaum Associates: Mahwah, NJ.

Wenger, M.J. & Townsend, J.T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 45,* 283-323.

Yang, C.-T. (2011). Relative saliency in change signals affects perceptual comparison and decision processes in change detection. *Journal of Experimental Psychology: Human Perception & Performance, 37*(6), 1708-1728.

Yang, C.-T., Chang, T.-Y., & Wu, C.-J. (2013). Relative change probability affects the decision process of detecting multiple feature changes. *Journal of Experimental Psychology: Human Perception & Performance, 39*(5), 1365-1385.

Yang, C-T. (2016). Attention and perceptual decision making. 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*. Elsevier.

Yang, H., Fifić, M., & Townsend, J. T. (2014). Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes.*Journal of Mathematical Psychology, 58*, 21-32.

Yu, J.-C., Chang, T.-Y., & Yang, C.-T. (2014). Individual differences in working memory capacity and workload capacity. *Frontiers in Psychology, 5*.

1. To use just one of many possible examples, tasting wine involves the integration of visual with olfactory cues. Interestingly, simply adding red food coloring to white wine has caused it to be described as a red wine by a panel of experts (Marrot, Brochet, & Dubourdieu, 2001). [↑](#footnote-ref-2)
2. For further information on historical precedents in this field, we refer to the reader to excellent reviews by Algom, Eidels, Hawkins, Jefferson, and Townsend (2015). [↑](#footnote-ref-3)
3. In this chapter, we use the term *channel* to refer to a sensory, cognitive, or psychological process directed toward a specific stimulus element or goal. [↑](#footnote-ref-4)