Psychological Review

ISSN: 0033-295X

https://doi.org/10.1037/rev0000232

Extending Systems Factorial Technology to Errored Responses

Daniel R. Little[1], Haiyuan Yang[2], Ami Eidels[3], and James T. Townsend[2]
[1] School of Psychological Sciences, The University of Melbourne
[2] Indiana University
[3] The University of Newcastle

Systems factorial technology (SFT) is a theoretically derived methodology that allows for strong inferences to be made about the underlying processing architecture (e.g., whether processing occurs in a pooled, coactive fashion or in serial or in parallel). Measures of mental architecture using SFT have been restricted to the use of error-free response times (RTs). In this article, through formal proofs and demonstrations, we extended the measure of architecture, the survivor interaction contrast (SIC), to RTs conditioned on whether they are correct or incorrect. We show that so long as an ordering relation (between stimulus conditions of different difficulty) is preserved, we learn that the canonical SIC predictions result when exhaustive processing is necessary and sufficient for a response. We further prove that this ordering relation holds for the popular Wiener diffusion model for both correct and error RTs but fails under some classes of a Poisson counter model, which affords a strong potential experimental test of the latter class versus the others. Our exploration also serves to point to the importance of detailed studies of how errors are made in perceptual and cognitive tasks.

Keywords: systems factorial technology, response times, nonparametric analyis, serial versus parallel systems, errors

There have been two general approaches to the study of human response times (RTs; Logan, 2002; Townsend et al., 2018). One method focuses on developing a parametric model capable of explaining important observed characteristics of RT data. Exemplars of this approach include the set of sequential sampling models such as Relative Judgment Theory (Link & Heath, 1975), Diffusion Decision Model (DDM; Ratcliff, 1978; Ratcliff et al., 2016), the Linear Ballistic Accumulator (LBA; Brown & Heathcote, 2008), and their extensions (e.g., the Leaky Competing Accumulator; Usher & McClelland, 2001) or competitors (e.g., Decision Field Theory, based on Ornstein–Uhlenbeck stochastics; Busemeyer & Townsend, 1993). Using psychologically plausible parameters, these models successfully explain the relation between choices and RTs, speed–accuracy trade-offs, the shapes of RT distributions, choice bias, the relation between correct and error RTs, and even classical findings and laws of psychophysics (Link, 1992; Luce, 1986; Ratcliff et al., 1999; Ratcliff & Rouder, 1998; Van Zandt & Ratcliff, 1995). This is useful because it allows decisions to be decomposed into psychologically meaningful parameters governing the rate of information accumulation or drift rate, how cautious or conservative responding is, and other nondecision processes. These models have also provided useful generalizations and approximations to neural firing rates (measured through single-cell recording, electroencephalography [EEG], functional magnetic resonance imaging [fMRI]), hence providing important bridging links, via latent model parameters, between behavior and the underlying neural mechanisms (Forstmann & Wagenmakers, 2015; Smith, 2010).

A key limitation of the single model approach is that it makes strong assumptions about the architecture of information processing, and in turn, additional assumptions about other aspects such as stopping rule and processing capacity. To wit, these assumptions have been recently challenged on several grounds (Jones & Dzhafarov, 2014, but see Smith et al., 2014 and Heathcote et al., 2014). Parameterized models bear other potential weaknesses. For instance, even models possessing the exact same number of parameters can differ greatly in their complexity and therefore their innate data-fitting ability. This fact can be witnessed by comparing two competing models of pattern recognition and confusion: The popular Similarity Choice Model (Luce, 1963) has been demonstrated to be significantly more complex than a natural competitor, the Overlap Model (Townsend, 1971b; see also Myung & Pitt, 2009; Townsend & Landon, 1983). Finally, one model may predict data better simply because of its specific distributional assumptions rather than because of more fundamental psychological characteristics.

The second approach focuses on developing theoretically motivated methods for deriving qualitatively distinct predictions from

Daniel R. Little https://orcid.org/0000-0003-3607-5525
This work was supported by ARC Discovery Project Grant DP160102360 to authors Daniel R. Little, Ami Eidels, and James T. Townsend.

The authors Daniel R. Little and Haiyuan Yang contributed equally to this article and share first authorship. Portions of this article were completed as part of a PhD Thesis by Haiyuan Yang. Portions of this work have been presented at the Annual Conference of the Society for Mathematical Psychology. A preprint of this work was published in H. Yang et al. (2020). This work is not preregistered. We thank Matt Jones for helpful comments on earlier drafts of this work.

Poisson model simulation code is available at: https://github.com/knowlab Unimelb/SFTERRORS.

Correspondence concerning this article should be addressed to Daniel R. Little, School of Psychological Sciences, The University of Melbourne, Melbourne, VIC 3010, Australia. Email: daniel.little@unimelb.edu.au

Template Version: 22 January 2022 ■ 6:35 pm IST REV-2020-1568_blupencil ■ 14 February 2022 ■ 2:24 pm IST

large classes of information processing models. These approaches often attempt to avoid assumptions about specific distributions of either model components or outcome variables, making them general and powerful. One of the major goals of such approaches is to design accompanying experimental methods that avoid model mimicry, a common specter in the behavioral sciences. Even two rigorously defined mathematical models based on critically distinctive assumptions can oftentimes mimic one another's predictions, even to the point of mathematical equivalence (Greeno & Steiner, 1964; Jones & Dzhafarov, 2014; Khodadadi & Townsend, 2015; Townsend, 1971a).

One meta-theoretic approach, systems factorial technology (SFT; Townsend et al., 2018; Townsend & Nozawa, 1995), has systematically built upon and generalized many of the key historical theories including, among others, the method of subtraction (Donders, 1868), additive factors methods (Schweickert, 1978; Sternberg, 1966, 1969; Townsend, 1971a, 1984), redundant target methods (Miller, 1982; Raab, 1962), and trichonometric theory (Schweickert & Townsend, 1989). SFT has been reviewed in several contemporary tutorials (Algom & Fitousi, 2016; Altieri et al., 2017; Harding et al., 2016; Townsend et al., 2018) in the recent volume by Little, Altieri, et al. (2017) and a special issue of the *Journal of Mathematical Psychology* (Houpt et al., 2019). The present article extends the inferences of SFT to included RTs conditioned on whether the response was correct or incorrect.

SFT has enabled important results in a number of important areas, but these results are limited to the analysis of correct RTs. The analytical tools and methods encompassed under the umbrella of SFT have led to advances in visual search (Fific, Townsend, et al., 2008), categorization (Blunden et al., 2015; Fific et al., 2010; Fific, Nosofsky, et al., 2008; Little et al., 2011, 2013; Moneer et al., 2016), recognition memory (Townsend & Fific, 2004), face processing (Cheng et al., 2018; Fific & Townsend, 2010), numerical processing (Garrett et al., 2019), change detection (Blunden et al., 2021; C.-T. Yang et al., 2011, 2013), long-term memory retrieval (Cox & Criss, 2017; Howard et al., 2020; Shang et al., 2021), single-unit firing patterns in macaques (Lowe et al., 2019), and simple detection (C.-T. Yang, Little, et al., 2014). Townsend et al. (2018) present a recent overview of SFT accompanied by an annotated list of applications. The underlying theorems supporting inferences in each of these cases and others are based on error-free responding.

As an example, consider an implementation of SFT to the study of categorization. Fific, Nosofsky, et al. (2008) were interested in the characteristics of processing that differentiate the categorization of integral dimension from separable dimension stimuli. Integral dimensions, like brightness and saturation, are difficult to attend in isolation; on the other hand, separable dimensions, such as shape and color, can be analyzed independently (Algom & Fitousi, 2016; Ashby et al., 1994; Burns, 2016; Garner, 1974; Nosofsky & Palmeri, 1997a; Shepard & Chang, 1963; Shepard et al., 1961). Using SFT, Fific et al. found that integral dimensions were best characterized by overadditivity at the distributional level; by contrast, separable dimensions were best characterized by additivity or underadditivity at the distributional level. Although there are many converging results demonstrating a difference between integral and separable dimensions (see Griffiths et al., 2017 for a review), this was a much stronger demonstration of a difference in dimensional processing than previously demonstrated because observation of additivity or underadditivity for separable dimension rules out a coactive processing architecture. Until that point, the leading categorization RT models (e.g., the exemplar-based random walk model; Nosofsky & Palmeri, 1997b) assumed coactivity. Follow-up studies have confirmed and extended this result for many different types of dimensions (Blunden et al., 2015, 2020; Cheng et al., 2018; Fific et al., 2010; Little et al., 2011, 2013; Moneer et al., 2016).

The appeal and theoretical importance of SFT lies in its ability to strongly differentiate several fundamental attributes of information processing systems (Algom et al., 2015; Townsend & Wenger, 2004). The implication is that formidable model mimicry challenges alluded to above are thereby obviated. For instance, are multiple sources of information processed in serial or in parallel? Does processing stop after processing one source or continue exhaustively through all sources? Are processes independent or are there facilitatory or inhibitory interactions between channels? And what is the capacity of the information processing system? This line of research has been particularly important to the continued development of methods that can explain both choices and RTs at the full distributional level while avoiding model mimicry between serial and parallel systems (among others).

While SFT has successfully engaged a multitude of issues in perception and cognition, the theory has been limited to highly accurate performance. That is, SFT, at least applied to questions of information processing architecture and stopping rule, does not address error rates or error RTs. Although evidence will be cited later on regarding the robustness of SFT predictions up to moderately high error rates, it is undeniable that extension of the theory and its implied methodology would be beneficial. To underscore the importance of this lacuna, error RTs, both mean values and their distributional shape, have provided a key source of evidence in the parametric model approach allowing specific instances of that model to be ruled out on the basis of empirical data. For instance, the Wiener diffusion model, a simpler model than the DDM that does not allow for between-trial variation in drift rate, cannot explain the pattern of error RTs found in many experiments. In fact, in a task without response bias, the Wiener diffusion model is limited to predicting error RTs which have the same mean as correct RTs (Diederich & Busemeyer, 2003), a prediction not typically observed in data.

It is arguable that RTs and response frequencies are the only observable behavioral variables that lie on strong measurement scales (Roberts, 1979; Townsend, 1992), and as a consequence, are intimately related to internal processing mechanisms and their dynamics. Increasing the application of SFT to explain both correct and error RTs adds powerful additional constraints to competing hypotheses and models over and above the focus on a single dependent variable. An unavoidable outcome of increasing the scope of the analysis to include additional observable variables is that the attendant models increase in complexity. This additional complexity affords a double-edged sword since the related behaviors emerge in a more intricate fashion and canonical experimental distinctions are sometimes limited. However, the ensuing diversity also offers new opportunities for assessing more detailed psychological issues connected to models and their predictions.

Brief Overview of SFT

Broadly speaking, there are three major threads to SFT, accompanied by appropriate diagnostic analyses: *capacity*, *resilience*, and *architecture*. The study of capacity makes use of the Capacity

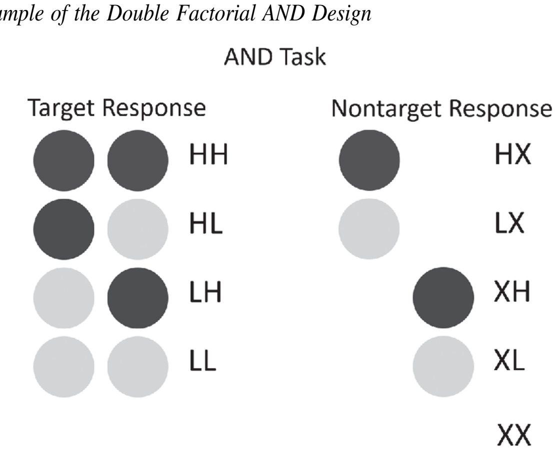
Coefficient, which informs researchers about the system's ability to cope with increased load (typically an increase in the number of signals presented for processing). The theory underlying the measurement of workload capacity has already been extended to error RTs, termed the Assessment Function (Donkin et al., 2014; Townsend & Altieri, 2012). The Assessment Function opens the methodological gateway to critical studies involving speed–accuracy trade-offs (Donkin et al., 2014). The study of resilience makes use of the Resilience and Conflict Contrast Functions, which assess how functions cope with distracting or conflicting information (Houpt & Little, 2017; Little et al., 2015, 2018). The form of the resilience function is highly similar to the capacity coefficient but adds additional contrasts between differing levels of distractor salience. Finally, the study of architecture uses the survivor interaction contrast (SIC), which informs whether processing is serial, parallel, or coactive and whether the stopping rule is exhaustive or self-terminating. To date, the essential theory for differentiating architecture and stopping rules, via the SIC, applies only to correct RTs. The purpose of this article is to provide new theorems extending the nonparametric SIC results to RTs conditioned on correct or error responding. In addition, we investigate some popular parameterized models in order to assess whether they meet the conditions important to the validity of the theorems as well as to exemplify those theorems.

In the following, we define the nonparametric forms of the conditional RT distributions (conditioned on correct or error responding) for different classes of model architecture. SFT makes use of the double factorial design (Townsend & Nozawa, 1995). Figures 1 and 2 show examples of a double factorial detection experimental design using an AND rule and an OR rule, respectively.

In the double factorial design, the observer is presented with a pair of stimuli at two locations. The observers task is to detect the presence of a target at either one or both locations (OR task) or at both locations (AND task). As indicated in Figures 1 and 2, the salience of the target varies but is irrelevant to the response.

Figure 1

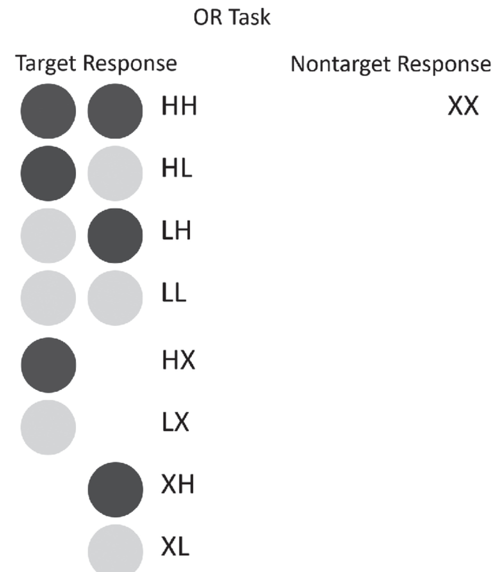
Example of the Double Factorial AND Design



Note. Target response stimuli have a factorial manipulation of target salience in two locations. Nontarget response stimuli are present in either one or neither location. Dark, light, and absent circles represent high salience (H), low salience (L), and null (X) target stimuli, respectively.

Figure 2

Example of the Double Factorial OR Design



Note. Target response stimuli have a target in either two or one location. Nontargets have no targets presented in either location. Dark, light, and absent circles represent high salience (H), low salience (L), and null (X) target stimuli, respectively.

Both high and low salience targets are treated as present. Varying salience, however, influences the RT with which a detection decision is made.

It is important to note that the experimental design often dictates the appropriate stopping rule. For instance, in an AND design (and by definition, therefore using a conjunctive rule), processing of all items, dimensions, and so forth of a designated target class must be completed before a "positive" response can be made. A "negative" response will be made if any of the dimensions or features are not members of that class. Here we define an *exhaustive* stopping rule as processing all available sources of information regardless of whether a decision could logically be made earlier. By contrast, a *self-terminating* stopping rule is one in which processing stops as soon as a response can be determined. In an AND design, exhaustive and self-terminating responding can be differentiated on the basis of negative, nontarget responses but not on the basis of positive, target responses. These definitions correspond to similar definitions used by Sternberg (1969) in his classic study of recognition memory.

On the other hand, in an OR design using a disjunctive rule, a positive response can be made if any of the dimensions or features is a member of the positive target class. In contrast, a negative response must be made only if none of the presented dimensions are of that class (i.e., they are *all* members of a *different* class, or classes, e.g., distractors or noise items). Hence, in an OR condition, the nontarget class requires processing of all items, dimensions, and so forth, thus preventing the differentiation of exhaustive and self-terminating stopping rules. By contrast, positive target responses in the OR task can be used to differentiate exhaustive and self-terminating processing. OR conditions are often found in so-called redundant signals experimental designs.

Inappropriate or inefficient stopping rules can be used and still lead to a correct outcome. For instance, while an exhaustive stopping rule is appropriate for target responses in an AND task, exhaustive stopping can also be used for target responses in an OR task and will still produce accurate, albeit temporally inefficient, responses. Conversely, application of inappropriate rules to certain scenarios can lead to errors, and errors can be generated from a system in a variety of ways. For instance, when a double target is presented in an AND task, a system will generate errors if it fails to detect a target in one location. The timing of the error will depend on whether the system continues to exhaustively process the other location or whether the system terminates processing at that point. In an OR task, exhaustive models can respond accurately, if both locations are identified appropriately but are again inefficient since processing can legitimately terminate after a target location is found.

Using the existing suite of SFT methods, in actual experiments using the double factorial paradigm (see, e.g., Little et al., 2011), in both AND and OR conditions, we often observe that participants adopt self-terminating processing for the disjunctive rule items (nontargets in the AND condition, see Figure 1; targets in the OR condition, see Figure 2) and exhaustive processing for the conjunctive rule items (targets in the AND condition, see Figure 1; nontargets in the OR condition, see Figure 2). A prevailing response pattern is that processing corresponds to a self-terminating system which follows the appropriate response rule: AND target responses follow from exhaustive processing but AND nontarget responses are the results of self-termination; OR target response are the results of self-terminating processing but nontarget responses follow from exhaustive processing. This type of response strategy allows for accurate responding to all items but also efficient RTs (see, e.g., Fific, Nosofsky, et al., 2008).

There are, naturally, many other kinds of self-terminating situations, a highly important one being single-target self-termination (e.g., Atkinson et al., 1969; Blaha & Houpt, 2015; Sternberg, 1966). Another important system is one which always terminates after the first process is complete (termed first-terminating or minimum time response). The latter two rules would violate the response rule in an AND task and consequently generate very high error rates. This type of responding has been termed rule-breaking for this reason (Bushmakin et al., 2017). Overall, these errors can be thought of as a consequence of misapplying (deliberate or accidental) an appropriate rule to the task at hand. Consequently, inefficient exhaustive processing in an OR task can also be considered as a type or rule-breaking. We do not focus on rule-breaking models in this article. Rather, our present focus is on errors of another kind, that are generated from what might be termed *accumulator failures*, in which the identification of a target or nontarget in either location fails because information arriving from processing that location was incorrect.

Another fundamental consideration is whether processing of both channels occurs in serial or in parallel. The architecture of the system can be completely crossed with the stopping rule, and we refer to Parallel Exhaustive, Parallel Self-Terminating, Serial Exhaustive, and Serial Self-Terminating models. As explained in the section Survivor Interaction Contrasts Conditioned on Accuracy and depending on the task, any of these models may result in error if the individual processing channels identify the wrong value of a feature or item. The timing of that error will depend on the combination of architecture and stopping rule.

The application of the SIC applies to the double-target items (i.e., those experimental displays where both signal locations are occupied by target items such as the the "target response stimuli" on the left-hand side of Figure 1). Errors occur when these double targets are mistakenly confused for nontarget items. In an OR task, double targets can also be confused for single targets but still result in a correct response. The timing of the correct and error responses depends on the architecture and the stopping rule. We direct the reader to Table 1 which enumerates the models (architecture plus stopping rule) that we consider along with task and response pattern to which it applies. We make a further descriptive distinction of the type of response pattern evident for the double targets. For example, while an Exhaustive model is always exhaustive—a Self-Terminating model may self-terminate depending on the task and whether the response is correct or in error.

One major challenge is that a key assumption to legitimate employment of SIC functions, namely *selective influence* at the distributional ordering level (Townsend, 1990b), is not always satisfied. We discuss the selective influence assumption in the section below. In addition, we will see that, depending on the stopping rule, the complexity of correct or error responding increases dramatically. Nonetheless, when the model circumstances are benign, we can utilize distributions that are conditional on being correct or incorrect in order to analyze the key nonparametric contrasts (conditional SIC). In such instances, we can demonstrate the consequent power of SIC functions in assessing architecture and stopping rule. In short, the diagnosticity of SFT holds at any level of error so long as a key assumption (about the ordering of the RT distributions across stimulus conditions) is met. We prove this to be true in the case of exhaustive processing conditional on correct responding for the AND task and show that this is the case using simulations for the self-terminating models conditional on correct responding.

Selective Influence and the Critical Statistical Function: SIC

To diagnose processing architecture in the conditions we study here, the SIC relies on analysis of the RTs of the double-target stimuli (see Figures 1 and 2). For these stimuli, there is a target in each location, and the difficulty of each of the target decisions has been manipulated. For example, in a detection task where a target is defined by a luminance contrast difference from background, these four stimulus combinations would be comprised of the following: (a) high salient targets in both A and B channels (we term this stimulus HH), (b) a high salient target in the A channel but a target of lower salience in the B channel (HL), (c) a low salient target in A and a high salient target in B (LH), and (d) a low salient target in both channels (LL). Of course, in such tasks, there are also single-target trials (with only one target), and trials with no target at all, as seen on the right-hand side of the two figures. These trials ensure participants must process the presented display before responding and are also useful for other analyses (such as resilience; Little et al., 2015) but not for calculations of the SIC.

We now briefly recount the logic attendant to how selective influence can be manifested in data with sufficient strength to identify architecture and stopping rules. Townsend and colleagues (Townsend, 1984, 1990b; Townsend & Ashby, 1978, 1983) investigated the dominance relationship between random variables and

Table 1
List of Accumulator Failure Models and the Task and Response to Which They Apply

Response	Task	Architecture	Stopping rule	Equation	Proposition	States
Correct	Conjunctive (AND)	Parallel	Exhaustive	5	1	1
Correct	Conjunctive (AND)	Parallel	Self-Terminating	6	2	1
Correct	Conjunctive (AND)	Serial	Exhaustive	8	3	1
Correct	Conjunctive (AND)	Serial	Self-Terminating	9	4	1
Correct	Disjunctive (OR)	Parallel	Exhaustive	5	[a]	1–3
Correct	Disjunctive (OR)	Parallel	Self-Terminating	13	[a]	1–3
Correct	Disjunctive (OR)	Serial	Exhaustive	8	[a]	1–3
Correct	Disjunctive (OR)	Serial	Self-Terminating	14	[a]	1–3
Error	Conjunctive (AND)	Parallel	Exhaustive	5	[a]	2–4
Error	Conjunctive (AND)	Parallel	Self-Terminating	6	[a]	2–4
Error	Conjunctive (AND)	Serial	Exhaustive	8	[a]	2–4
Error	Conjunctive (AND)	Serial	Self-Terminating	9	[a]	2–4
Error	Disjunctive (OR)	Parallel	Exhaustive	5	1[b]	4
Error	Disjunctive (OR)	Parallel	Self-Terminating	13	2[b]	4
Error	Disjunctive (OR)	Serial	Exhaustive	8	3[b]	4
Error	Disjunctive (OR)	Serial	Self-Terminating	14	4[b]	4

Note. "States" refer to the states of the model accumulators, shown as lines in each model equation, that enter into the SIC for each model. Formal definitions are provided at the equations listed along with the propositions for the conditional SICs. SIC = survivor interaction contrast.
[a] The SIC is comprised of a mixture of states where both channels are correct (or in error) or where only one channel is correct (or in error). We rely on simulations for these models. [b] The roles of the correct and error accumulators in these equations should be swapped.

proposed several measures that define a hierarchy of implication. Readers already versed in these matters can skip to the next section.

We introduce four measures that are most relevant for our purposes. The first two measures are the mean and the cumulative distribution function (CDF). Suppose there are two random variables $T1$ and $T2$. If the CDF of $T1$ is larger than the CDF of $T2$, that is, $F1(t) > F2(t)$ for all t, we say $T1$ and $T2$ are stochastically ordered, or $T1 < T2$; we also refer to this situation as the stochastic dominance of $F1(t)$ over $F2(t)$. This ordering of distribution functions implies that the mean of the random variables is also ordered, that is, $E(T1) < E(T2)$. In addition, since survivor functions $S(t)$ are defined as one minus the CDFs, it is clear that an ordering of the CDFs is equivalent to an ordering of the survivor functions.

The third measure is the hazard function, which measures the ordering between two variables at a stronger level. The hazard function is given by dividing the density function by the survivor function: $h(t) = f(t)/S(t)$. Townsend and Ashby (1978) showed that an ordering of hazard functions, $h1(t) > h2(t)$ for all t, is stronger than an ordering of the associated CDFs, $F1(t) > F2(t)$ for all t, in the sense that the former implies the latter but not vice versa. The proof is based on the relationship between the hazard function and the survivor function: $-ln[S(t)] = \int_0^t h(s)ds$.

The final measure involves the likelihood ratio function, $L(t)$. For random variables, T_1 and T_2, we define the likelihood ratio function as $L(t) = f_1(t)/f_2(t)$. If $L(t)$ is monotonically decreasing with t, then $T1$ and $T2$ are ordered at the likelihood ratio level, which implies orderings at both the hazard function level and the CDFs.

Note that channels A and B might entail differing overall levels of difficulty such that the RT for HL does not equal the RT for LH. Such equality is unnecessary. This factorial manipulation of difficulty on each channel allows one to construct a set of contrasts that can be used to identify distinct processing architectures. For instance, SFT generalizes the additive factors method by showing that although serial models predict additivity at the mean RT level, as Sternberg showed for either self-termination or exhaustive processing (Sternberg, 1969), parallel models predict underadditivity or overadditivity depending on whether the stopping rule is exhaustive or self-terminating, respectively (Schweickert, 1978; Schweickert & Townsend, 1989; Townsend, 1984; Townsend & Nozawa, 1995). This test can be efficiently summarized in the mean interaction contrast (MIC for short), with $\bar{M}$, being the mean RT from the four factorial conditions:

$$\text{MIC} = [\bar{M}_{\text{LL}} - \bar{M}_{\text{LH}}] - [\bar{M}_{\text{HL}} - \bar{M}_{\text{HH}}]. \tag{1}$$

A much stronger test arises through contrasting the distributions for each of the factorial targets (Sternberg, 1969; Townsend, 1990b). The measure is based on the interaction of the survivor functions. It is helpful here to highlight the notation that we use throughout. $f(t)$ refers to the probability density function (PDF). In the context of speeded decisions, the PDF is the probability that processing had terminated (and a response is given) at time t. The PDF can be estimated empirically using the histograms of RTs from each condition (van Zandt, 2000). $F(t) = \int_0^t f(t)dt = P(T \le t)$ refers to the CDF and gives the probability that a positive random variable time T is less than or equal to some time t. The survivor function, $S(t) = 1 - F(t) = \int_t^\infty f(t)dt = P(T > t)$, gives the probability that a positive random variable time T is greater than some time t; that is, the probability that the process has not ended by time t. The SIC is then given as:

$$\begin{aligned}\text{SIC}(t) &= [S_{\text{LL}}(t) - S_{\text{LH}}(t)] - [S_{\text{HL}}(t) - S_{\text{HH}}(t)] \\ &= [P_{\text{LL}}(T > t) - P_{\text{LH}}(T > t)] - [P_{\text{HL}}(T > t) - P_{\text{HH}}(T > t)].\end{aligned} \tag{2}$$

Townsend and Nozawa (1995) showed that several classes of architecture and stopping rule make qualitatively different nonparametric predictions for the SIC. Under some minimal assumptions, outlined below, the different models cannot mimic each other, which breaks a long-standing limitation of testing these models

(Little, Eidels, et al., 2017; Townsend, 1990a). Consequently, estimating the MIC and SIC from empirical data in the four factorial conditions (HH, HL, etc.) can be used to rule out whole model classes. Figure 3 shows the SIC shapes for several processing models. H. Yang, Fific, et al. (2014) generalize these signatures to an arbitrary number of processes (i.e., beyond two). In the next section, we derive proofs for SIC signatures where performance is imperfect, conditioned on correct and error responses.

SICs Conditioned on Accuracy

Another arm of SFT, which we only briefly mention here, concerns with the processing capacity of the system. Townsend and colleagues (Townsend & Nozawa, 1995; Townsend & Wenger, 2004) demonstrated how different models produce different capacity functions that capture the system's ability to handle changes in the amount of information presented for processing. These seminal studies focused on high accuracy, error-free performance. In order to derive the Assessment Function, thereby generalizing the theory of capacity to account for correct and error RTs, Townsend and Altieri (2012) considered the properties of a pair of accrual halting models (i.e., a pair of accumulators racing to a threshold; see Figure 4). Each pair of accumulators represents a single-target channel (e.g., A or B), and each can either terminate correctly or in error. For simplicity, we refer to correct and incorrect accumulators in each channel rather than the actual stimulus values detected in each channel. For example, in an AND task, a double target will elicit an incorrect response if either Channel A or Channel B (or both) finish in error (where error in this case is the incorrect determination that a target signal was not present). That is, in an AND task, the response for a redundant target will be correct only if a target is detected correctly in both channels; hence, should any channel terminate incorrectly, the overall response will be in error. As explained below, the timing of this error will depend on whether the stopping rule is exhaustive or self-terminating and whether the architecture is parallel or serial.

The completion time for a pair of accumulators within a channel (e.g., T_{a1} and T_{a2}) is given by the completion time of the first accumulator to finish (e.g., the completion time is equal to the T_{a1} completion time if $T_{a1} \leq T_{a2}$). More formally, the overall completion time for Channel A is $T_a = \min(T_{a1}, T_{a2})$, where T_{a1} and T_{a2} is the completion time of each accumulator. Likewise, $T_b = \min(T_{b1}, T_{b2})$. The total RT, T, is a function of the finishing time on each channel: $T = f(T_a, T_b)$, where the form of f depends on how the two channels operate in the system (but not on the assumption of parallel processing within channels).

In the original formulation of the SIC, where each channel is always assumed to be correct, the RT is determined only by the architecture and the stopping rule. Specifically, in a Serial Exhaustive architecture, for a double target, the total RT is the sum of the completion times on each channel, $T^{\text{SerialEX}} = T_a + T_b$. In a Parallel Exhaustive model, the total RT is the maximum time for processing any single channel, $T^{\text{ParallelEX}} = \max(T_a, T_b)$. In a Self-Terminating model in an OR design, the system can terminate as soon as any channel completes correctly. Hence, the total RT in a Serial Self-Terminating model would equal $T^{\text{SerialST}} = T_a$ if Channel A is processed first and $T^{\text{SerialST}} = T_b$ if Channel B is processed first. In an OR task, for a Parallel Self-Terminating model, the total processing time is the minimum time to process either channel, $T^{\text{ParallelST}} = \min(T_a, T_b)$. Corresponding considerations are made for determining the RT distributions, f.

When accounting for error RTs, the RT depends not only on the architecture and stopping rule but also on the accuracy of each channel. For example, for a Parallel Self-Terminating model in an OR task with both targets present, the total RT is not necessarily the minimum of the two finishing times because when a target channel terminates incorrectly, the system may still produce a correct response if the other channel finishes correctly. In the case where Channel A finishes incorrectly (i.e., "no target signal in A") but Channel B finished correctly ("target present in B"), and the overall response is correct, the total RT is equal to T_b instead of $\min(T_a, T_b)$. On the other hand, if the overall response is incorrect, then the Parallel Self-Terminating model in the same condition would generate an RT equal to $\max(T_a, T_b)$ because the system would have to wait to confirm that the second channel also does not contain a target. Similar considerations apply to Parallel Exhaustive models and Serial models.

Conditional SIC Proofs

The conditional SIC predictions rely first on the assumption that the processing time for Channel A is independent of the processing time of Channel B, and second that the manipulation of salience on a channel has an effect only on that channel. For example, a high salience stimulus in Channel A results in a shorter correct finishing time on Channel A than a low salience stimulus on Channel A, regardless of what is happening on Channel B. Note that this assumption may be violated for errors for some reasonable models of RT which we review below. A third assumption is that of stochastic

Figure 3

Error-Free Survivor Interaction Contrasts for Each of the Canonical Processing Architectures

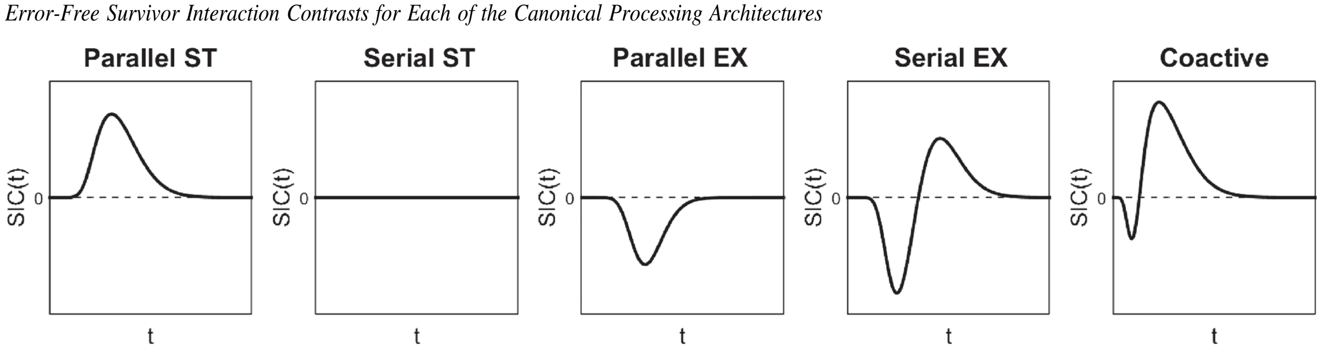
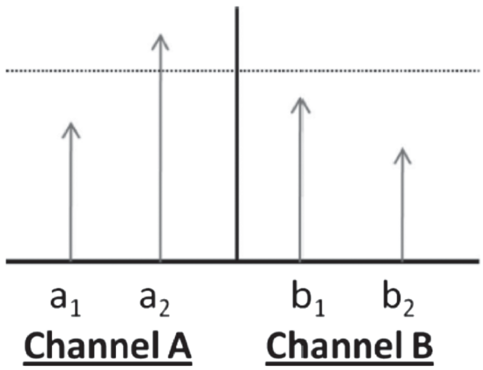


Figure 4
A Pair of Accrual Halting Models for Channel A and Channel B



Note. The first accumulator in each pair represents the correct processing of the attribute in that channel; the second accumulator in each pair represents an error.

dominance, which we elucidate further below. The original error-free proofs in Townsend and Nozawa (1995) did not assume independence. They allowed dependence among the processes but imposed additional conditions that were sufficient to garner the result. Here, we assume stochastic independence for simplicity along with the important condition of stochastic dominance at the distributional level (Townsend, 1990b). The error-free proofs are stronger in this regard because stochastic independence in processing is likely violated by global cognitive variables, such as attention.

Next, recall that in the error-free methodology, higher salience always speeds up a decision. In the case of less accurate functioning, we might wish to make the same assumption. As we show in our simulations, for correct responses, this tenet holds true. However, we shall see that some models predict faster errors on a low salience setting than on a high setting. For now, to aid intuition, we shall simply consider correct responses with the postulate that salience speeds up decisions in the sense that

$$P_H(T \le t|\text{CR}) > P_L(T \le t|\text{CR}). \tag{3}$$

We will discover that this holds nicely for all our models but again, we find intriguing diversity for incorrect responses.

Because the conditional probabilities constitute a true probability space that sums to one [e.g., $P_H(T \le t|\text{CR}) + P_H(T > t|\text{CR}) = 1$], an ordering of correct response probabilities implies the converse ordering on survivor functions. The SIC conditioned on accuracy can then be written as:

$$\text{SIC}_{\text{correct}}(t) = [S_{\text{LL}}(t|\text{CR}) - S_{\text{LH}}(t|\text{CR})] - [S_{\text{HL}}(t|\text{CR}) - S_{\text{HH}}(t|\text{CR})]. \tag{4}$$

It is clear that the $\text{SIC}_{\text{correct}}(t)$ is a generalization of the error-free SIC developed by Townsend and Nozawa (1995); when accuracy approaches 1, $\text{SIC}_{\text{correct}}(t) = \text{SIC}(t)$. In the following, we derive the conditional SIC predictions for Parallel Exhaustive, Serial Exhaustive, Parallel Self-Terminating, and Serial Self-Terminating models assuming that errors arise from a failure to identify a target in one or both of the channels.

Correct Responses

AND Task

Correct double-target trials in an AND task are straightforward since a model can only terminate correctly if both channels are correct. Error responses in the AND task are complicated by the fact that an error will be generated if either or both of the channels fail to detect a target. The added complexity means that stochastic dominance cannot be employed here for SIC predictions conditioned on error responding; yet it is necessary for theorists to enter this territory. The various ways of being correct or incorrect are a psychological reality and must ultimately be confronted, experimented on, and modeled. Since the complexities arise in a similar manner for each model, we present the formal definitions of the models first followed by the SICs.

Parallel Exhaustive Models

In an AND task, when a double-target stimulus is presented, both target channels need to be detected correctly in order to trigger a correct response. Our modeling approach assumes that processing of each target signal (or channel) happens on a separate channel, and within each channel, there is an accumulator for correct response and an accumulator for an incorrect response (as illustrated in Figure 4). A correct response is made only when the correct accumulators in each channel terminate before the incorrect accumulators in each channel. If either of the incorrect accumulators terminates first, then an error response would be made. This will occur because, in an AND task, the error channel provides evidence for a nontarget response. There are therefore four states depending on which accumulators finish first in each channel. A Parallel Exhaustive model predicts an exhaustive RT for each of these states:

$$\begin{aligned}
&\text{if } T_{\text{AC}} < T_{\text{AI}} \text{ and } T_{\text{BC}} < T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelEX}} = \max(T_{\text{AC}}, T_{\text{BC}})\\
&\text{if } T_{\text{AC}} < T_{\text{AI}} \text{ and } T_{\text{BC}} > T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelEX}} = \max(T_{\text{AC}}, T_{\text{BI}})\\
&\text{if } T_{\text{AC}} > T_{\text{AI}} \text{ and } T_{\text{BC}} < T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelEX}} = \max(T_{\text{AI}}, T_{\text{BC}})\\
&\text{if } T_{\text{AC}} > T_{\text{AI}} \text{ and } T_{\text{BC}} > T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelEX}} = \max(T_{\text{AI}}, T_{\text{BI}}).
\end{aligned} \tag{5}$$

Parallel Self-Terminating Models

For a Self-Terminating model in an AND task, when a double-target stimulus is presented, both target channels still need to be detected correctly in order to trigger a correct response. If either of the incorrect accumulators terminates first, then an error response would be made when the error channel triggers self-termination. There are four states depending on which accumulators finish first in each channel. A Parallel Self-terminating model predicts a different RT calculation for each of these states:

$$\begin{aligned}
&\text{if } T_{\text{AC}} < T_{\text{AI}} \text{ and } T_{\text{BC}} < T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelST}} = \max(T_{\text{AC}}, T_{\text{BC}})\\
&\text{if } T_{\text{AC}} < T_{\text{AI}} \text{ and } T_{\text{BC}} > T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelST}} = T_{\text{BI}}\\
&\text{if } T_{\text{AC}} > T_{\text{AI}} \text{ and } T_{\text{BC}} < T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelST}} = T_{\text{AI}}\\
&\text{if } T_{\text{AC}} > T_{\text{AI}} \text{ and } T_{\text{BC}} > T_{\text{BI}}, \text{ then } T_{\text{AND}}^{\text{ParallelST}} = \min(T_{\text{AI}}, T_{\text{BI}}).
\end{aligned} \tag{6}$$

Correct SICs: AND Task—Parallel Models

We first consider the SIC conditioned on the response being correct. For both the Parallel Exhaustive model (Equation 6) and the Parallel Self-Terminating model (Equation 1), the first state generates a correct response, but the remaining three generate error responses. Furthermore, in both cases, the correct RT is given by the maximum RT of both accumulators; hence, the conditional correct SIC proof proceeds identically for both cases.

Proposition 1. Assuming independence, selective influence, and stochastic dominance, the Parallel Exhaustive model, in an AND task, predicts that the conditionally correct SIC is negative for all t.

Proof. For convenience, we use $\Delta^2_{A,B}$ to denote the operation of the double difference on A and B. That is:

$$\Delta^2_{A,B} S_{AB}(t) = \Delta^2_{A=[L,H],B=[L,H]} S_{AB}(t) = (S_{\mathrm{LL}}(t) - S_{\mathrm{LH}}(t)) - (S_{\mathrm{HL}}(t) - S_{\mathrm{HH}}(t)). \quad (7)$$

Thus, the conditional SIC can be written as:

$$\mathrm{SIC}_{\mathrm{correct}}(t) = \Delta^2_{A,B} S_{AB}(t|\mathrm{CR}).$$

A graphical illustration of the application of the double difference function is shown in Figure 5. Equation 5 shows that a correct response is produced only when both channels are correct. Hence,

$$\mathrm{SIC}_{\mathrm{correct}}(t) = \Delta^2_{A,B}[S(\max(T_{\mathrm{AC}}, T_{\mathrm{BC}})|T_{\mathrm{AC}} < T_{\mathrm{AI}}, T_{\mathrm{BC}} < T_{\mathrm{BI}})].$$

Note that the max time can be computed as the product of the CDFs because A and B are independent; hence, $F^{\max}_{AB}(t|CR) = F_A(t|CR) \times F_B(t|CR)$. Consequently, the SIC can be computed as:

$$\begin{aligned}\mathrm{SIC}_{\mathrm{correct}}(t) &= \Delta^2_{A,B}(1 - [F_A(t|\mathrm{CR}) \times F_B(t|\mathrm{CR})]) \\ &= -F_{A=L}(t|\mathrm{CR}) \times F_{B=L}(t|\mathrm{CR}) \\ &\quad + F_{A=L}(t|\mathrm{CR}) \times F_{B=H}(t|\mathrm{CR}) \\ &\quad + F_{A=H}(t|\mathrm{CR}) \times F_{B=L}(t|\mathrm{CR}) \\ &\quad - F_{A=H}(t|\mathrm{CR}) \times F_{B=H}(t|\mathrm{CR}) \\ &= -(F_{A=L}(t|\mathrm{CR}) - F_{A=H}(t|\mathrm{CR})) \\ &\quad \times (F_{B=L}(t|\mathrm{CR}) - F_{B=H}(t|\mathrm{CR})).\end{aligned}$$

Due to the assumption of stochastic dominance, $F_L(t|\mathrm{CR}) < F_H(t|\mathrm{CR})$ for all t implying that the two factors in the product are negative. Hence, the product is positive but negated, which proves that the conditional SIC for the Parallel Exhaustive model is negative for all t like its error-free counterpart. □

Proposition 2. Assuming independence and selective influence, the Parallel Self-Terminating model, in an AND task, predicts that the conditionally correct SIC is negative for all t.

Proof. The proof is the same as for the Parallel Exhaustive model. □

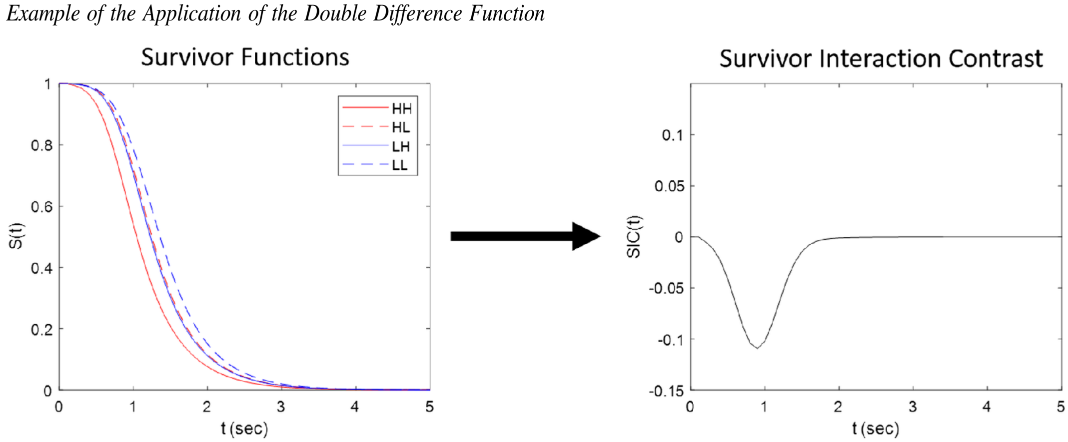
This proof shows that when conditioned on accuracy, the Parallel Exhaustive model and Parallel Self-Terminating model predict completely negative SICs in an AND task. Below, in the section Computation of Conditional SIC Functions, we show via computation that the shape of this SIC is commensurate with the shape of the error-free SIC prediction shown in Figure 3 for the Poisson model simulations considered in this article.

Serial Exhaustive Model

For a Serial Exhaustive model, in an AND task, the RT is always predicted as the sum of both processing channels:

Figure 5

Example of the Application of the Double Difference Function



$$SIC(t) = \Delta^2_{A,B}\left[S_{AB}(t)\right] = \left(S_{LL}(t) - S_{LH}(t)\right) - \left(S_{HL}(t) - S_{HH}(t)\right)$$

Note. Left panel: Survivor functions. Right panel: SIC function. Note that the survivor functions were sampled from a Parallel Exhaustive model when both channels are correct. SIC = survivor interaction contrast. See the online article for the color version of this figure.

$$\begin{aligned}&\text{if } T_{AC} < T_{AI} \text{ and } T_{BC} < T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AC} + T_{BC}\\ &\text{if } T_{AC} < T_{AI} \text{ and } T_{BC} > T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AC} + T_{BI}\\ &\text{if } T_{AC} > T_{AI} \text{ and } T_{BC} < T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AI} + T_{BC}\\ &\text{if } T_{AC} > T_{AI} \text{ and } T_{BC} > T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AI} + T_{BI}. \end{aligned} \quad (8)$$

Serial Self-Terminating Models

For a Serial Self-Terminating model, in an AND task assuming that Channel A is processed first, the processing time for the correct response (the first state) will be the same as the Serial Exhaustive model. The RTs for the error states, however, will differ depending on when the model can self-terminate. In the original SFT work, it is assumed that Channel A is processed first with a fixed probability, p, and that Channel B is processed first with probability $1 - p$. To simplify the case, we assume that Channel A is always processed first:

$$\begin{aligned}&\text{if } T_{AC} < T_{AI} \text{ and } T_{BC} < T_{BI}, \text{ then } T_{AND}^{SerialST} = T_{AC} + T_{BC}\\ &\text{if } T_{AC} < T_{AI} \text{ and } T_{BC} > T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AC} + T_{BI}\\ &\text{if } T_{AC} > T_{AI} \text{ and } T_{BC} < T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AI}\\ &\text{if } T_{AC} > T_{AI} \text{ and } T_{BC} > T_{BI}, \text{ then } T_{AND}^{SerialEX} = T_{AI}. \end{aligned} \quad (9)$$

Correct SICs: AND Task—Serial Models

For both the Serial Exhaustive and Serial Self-Terminating models, only the first state produces a correct response, the other states generate error responses.

Proposition 3. Assuming independence, selective influence, and stochastic dominance, the Serial Exhaustive model in an AND task predicts that the conditional SIC for correct responses is negative for small values of t and later becomes positive. Furthermore, the integral over the positive real line (i.e., of t) of the SIC function is 0.

Proof. The conditional SIC for the Serial Exhaustive model in an AND task has the following expression:

$$SIC_{correct}(t) = \Delta^2_{A,B}[S(T_{AC} + T_{BC}|T_{AC} < T_{AI}, T_{BC} < T_{BI})].$$

The survivor function for the sum of two variables is the sum of the probabilities that Channel A does not finish by time t and is correct, $S_A(t|T_{AC} < T_{AI})$, plus the probability that Channel A finishes before time t, but in the remaining time, Channel B does not finish; the latter term being the convolution of the density for the Channel A correct accumulator and the survivor function for the Channel B correct accumulator, $[f_A * S_B](t|T_{AC} < T_{AI}, T_{BC} < T_{BI})$. Hence, the conditional SIC can be written as follows:

$$SIC_{correct}(t) = \Delta^2_{A,B}\Big(S_A(t|T_{AC} < T_{AI}) + \int_0^t f_A(\tau|T_{AC} < T_{AI}) \times S_B(t-\tau|T_{AC} < T_{AI}, T_{BC} < T_{BI})d\tau\Big). \quad (10)$$

The first term cancels out in the double difference; expanding the second term gives:

$$\begin{aligned}SIC_{correct}(t) = &\int_0^t f_{A=L}(\tau|T_{AC} < T_{AI})S_{B=L}(t-\tau|T_{BC} < T_{BI})d\tau\\ &- \int_0^t f_{A=L}(\tau|T_{AC} < T_{AI})S_{B=H}(t-\tau|T_{BC} < T_{BI})d\tau\\ &- \int_0^t f_{A=H}(\tau|T_{AC} < T_{AI})S_{B=L}(t-\tau|T_{BC} < T_{BI})d\tau\\ &+ \int_0^t f_{A=H}(\tau|T_{AC} < T_{AI})S_{B=H}(t-\tau|T_{BC} < T_{BI})d\tau. \end{aligned} \quad (11)$$

Factoring the double difference function gives:

$$\begin{aligned}SIC_{correct}(t) = &\int_0^t [S_{B=L}(t-\tau|T_{BC} < T_{BI}) - S_{B=H}(t-\tau|T_{BC} < T_{BI})]\\ &\times [f_{A=L}(\tau|T_{AC} < T_{AI}) - f_{A=H}(\tau|T_{AC} < T_{AI})]d\tau. \end{aligned} \quad (12)$$

This equation has the same form as Theorem 4 in Townsend and Nozawa (1995; see also Proposition 2.1 in H. Yang, Fific, et al., 2014); our proof therefore follows the same logic: Due to the stochastic dominance assumption, $S_L(t|CR) > S_H(t|CR)$, and the first term is positive for all t.

For the second difference, $[f_{A=L}(\tau|T_{AC} < T_{AI}) - f_{A=H}(\tau|T_{AC} < T_{AI})]$ will be nonpositive for some τ less than t_* (and strictly negative for some τ) after which the difference will be positive. The conditional SIC for the Serial Exhaustive model is therefore negative for the initial component of the function, for $t \le t_*$ and then positive thereafter. Since the integral of the SIC over the real line equals 0, the positive integral will equal negative integral. If the densities cross more than once, then the SIC will cross 0 an odd number of times and $f_L(\tau)$ will be less than $f_H(\tau)$ for small t (Townsend, 1990b; Townsend & Nozawa, 1995; H. Yang, Fific, et al., 2014).

Finally, for any nonnegative random variable, T, $\int_0^\infty P(T > t)dt = E[T]$, where $E[T]$ is the expected value. Consequently, the integral of the SIC is equal to $E[T_{A=L} + T_{B=L}] - E[T_{A=L} + T_{B=H}] - E[T_{A=H} + T_{B=L}] + E[T_{A=H} + T_{B=H}]$ which equals 0. (This is evident since the expected value of the sum of two random variables is equal to the sum of the expected values of each random variable.) □

Proposition 4. Assuming independence, selective influence, and stochastic dominance, the Serial Self-Terminating model, in an AND task, predicts that the conditional SIC for correct responses is negative for small values of t and later becomes positive. The integral over the positive real line (i.e., of t) of the SIC function is 0.

Proof. The proof is the same as for the Serial Exhaustive model. □

OR Task

In an OR task, correct responses in models in which errors arise due to accumulator failures are more complicated in the sense that correct responses are sometimes made by "partially" correct judgments. That is, subjects only need to be correct on one channel to give a correct response. Because of this, stochastic dominance cannot be assured for the conditional correct SIC predictions. Given a pair of independent accumulators, for each model, there are again four possible states; however, in the OR design, the first three states correspond to correct responses while only the last corresponds to an error response.

Parallel Exhaustive Models

The RTs associated with each state of a Parallel Exhaustive model in an OR task are the same as in Equation 5. We list the states of the model here because in an OR design, the first three states lead to a correct response, while the last state leads to an error response.

$$\begin{aligned}
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelEX}} = \max(T_{\mathrm{AC}}, T_{\mathrm{BC}})\\
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelEX}} = \max(T_{\mathrm{AC}}, T_{\mathrm{BI}})\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelEX}} = \max(T_{\mathrm{AI}}, T_{\mathrm{BC}})\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelEX}} = \max(T_{\mathrm{AI}}, T_{\mathrm{BI}}). \quad (13)
\end{aligned}$$

Parallel Self-Terminating Models

An OR design permits systems which can self-terminate to evince a minimum time stopping operation. The RTs associated with each state of a Parallel Self-Terminating model are as follows:

$$\begin{aligned}
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelST}} = \min(T_{\mathrm{AC}}, T_{\mathrm{BC}})\\
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelST}} = T_{\mathrm{AC}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelST}} = T_{\mathrm{BC}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{ParallelST}} = \max(T_{\mathrm{AI}}, T_{\mathrm{BI}}). \quad (14)
\end{aligned}$$

Correct SICs: OR Task—Parallel Models

Note that first three states of both parallel models in the OR task generate correct responses. We denote the probability and the event of "Channel A being correctly processed" as $p_A = P(T_{AC} \leq T_{AI})$ and the probability and the event of "B being correctly processed" as $p_B = P(T_{BC} \leq T_{BI})$. The probability of A being processed incorrectly is $(1 - p_A)$. To simplify the notation, we refer to the event of Channel A and Channel B being incorrect as $\neg A$ and $\neg B$, respectively. The total probability of responding correctly is $P(\mathrm{CR}) = p_A p_B + (1 - p_A) p_B + p_A(1 - p_B)$. Hence, according to the law of total probability, the SIC conditioned on correct responding in the OR task is given as:

$$\begin{aligned}
\mathrm{SIC}_{\mathrm{correct}}(t) &= \Delta^2_{A,B} P(T > t|\mathrm{CR})\\
&= \Delta^2_{A,B}\left[P(T > t \cap \mathrm{CR}) \times \frac{1}{P(\mathrm{CR})}\right]\\
&= \Delta^2_{A,B}\left[P(T > t \cap [(A,B) \oplus (\neg A, B) \oplus (A, \neg B)]) \times \frac{1}{P(\mathrm{CR})}\right]\\
&= \Delta^2_{A,B}[P(T > t \cap (A,B)) + P(T > t \cap (\neg A, B))\\
&\quad + P(T > t \cap (A, \neg B))] \times \frac{1}{P(\mathrm{CR})}\\
&= \Delta^2_{A,B}[P(T > t|A,B) \times p_A p_B\\
&\quad + P(T > t|\neg A, B) \times (1 - p_A) p_B\\
&\quad + P(T > t|A, \neg B) \times p_A(1 - p_B)] \times \frac{1}{P(\mathrm{CR})}.
\end{aligned}$$

At this point, we are confronted with the problem that we cannot observe, nor presume on the basis of any of our assumptions, the probability that the system is in any of these correct states. Furthermore, these state probabilities are likely to differ in unknown ways between the LL, LH, HL, and HH stimuli. It is likely that accuracy rates in L channels are lower than H channels, but without this information, we cannot make a clear proof of the shape of the conditional SIC. In the General Discussion section, we highlight a new method which may allow identifiability of these states; for the present purposes, we show below (in our section: Computation of Conditional SIC Functions for Specific Distributions) the form of the SIC and examine the qualitative shape via simulation.

Serial Exhaustive Models

For a Serial Exhaustive model, the RTs for each state of the accumulators are given as the sum of the accumulator finishing times. Hence, the RTs for the Serial Exhaustive model in an OR task are the same as the Serial Exhaustive model in the AND task (see Equation 8). However, the models make different responses with the first three states leading to correct responses and the last state leading to an error.

$$\begin{aligned}
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialEX}} = T_{\mathrm{AC}} + T_{\mathrm{BC}}\\
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialEX}} = T_{\mathrm{AC}} + T_{\mathrm{BI}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialEX}} = T_{\mathrm{AI}} + T_{\mathrm{BC}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialEX}} = T_{\mathrm{AI}} + T_{\mathrm{BI}}. \quad (15)
\end{aligned}$$

Serial Self-Terminating Models

Serial Self-Terminating models imply that the two channels are processed sequentially, and the process is finished whenever a target is detected, which means that when both targets are presented, subjects may provide a correct response even if an error is made in processing the first channel. In the following, we give the expression of the RT conditioned on the judgment of the two channels. We again assume that Channel A is always processed first.

$$\begin{aligned}
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialST}} = T_{\mathrm{AC}}\\
&\text{if } T_{\mathrm{AC}} < T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialST}} = T_{\mathrm{AC}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} < T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialST}} = T_{\mathrm{AI}} + T_{\mathrm{BC}}\\
&\text{if } T_{\mathrm{AC}} > T_{\mathrm{AI}} \text{ and } T_{\mathrm{BC}} > T_{\mathrm{BI}}, \text{ then } T_{\mathrm{OR}}^{\mathrm{SerialST}} = T_{\mathrm{AI}} + T_{\mathrm{BI}}. \quad (16)
\end{aligned}$$

Correct SICs: OR Task—Serial Models

For both Serial models, the first three states result in correct responses, and the final state is an error response. Denote $P(T_{\mathrm{AC}} < T_{\mathrm{AI}})$ as p_A and $P(T_{\mathrm{BC}} < T_{\mathrm{BI}})$ as p_B. The total probability of responding correctly (CR) is $P(\mathrm{CR}) = p_A + (1 - p_A)p_B$. Thus, we can write the conditional SIC function as:

$$\begin{aligned}
\mathrm{SIC}_{\mathrm{correct}}(t) &= \Delta^2_{A,B} P(T > t|\mathrm{CR})\\
&= \Delta^2_{A,B}\left[P(T > t \cap \mathrm{CR}) \times \frac{1}{P(\mathrm{CR})}\right]\\
&= \Delta^2_{A,B}\left[P(T > t \cap [A \oplus (\neg A, B)]) \times \frac{1}{P(\mathrm{CR})}\right]\\
&= \Delta^2_{A,B}[P(T > t \cap A) + P(T > t \cap (\neg A, B))] \times \frac{1}{P(\mathrm{CR})}\\
&= \Delta^2_{A,B}[P(T > t|A) \times p_A + P(T > t|\neg A, B) \times (1 - p_A) p_B]\\
&\quad \times \frac{1}{P(\mathrm{CR})}.
\end{aligned}$$

We are again left with a conditional SIC function which is comprised of a mixture of states and rely on simulations to examine these SICs.

Interim Summary I

We provided proofs for the SICs conditioned on correct responses and found that, for the AND task, these proofs mirror the error-free SIC proofs (Townsend & Nozawa, 1995). For the OR task, the SICs conditioned on correct responses reflect a mixture of possible correct states. This mixture prevents us from determining the appropriate dominance relations, which in turn prevents establishing a proof of the SIC. In the next section, we will outline the SICs for the error responses and find a similar relationship.

Error Responses

AND Task

In an AND task, errors are generated whenever any of the channels terminates at a nontarget response. Observe that this symmetry (between errors in the AND task and correct responses in the OR task) means that similar considerations apply.

Parallel Models

For a parallel model, the last three states of Equations 5 and 6 all lead to error responses, the total error RT is a mixture of these expressions. According to the law of total probability, we can express the error conditioned SIC as:

$$\begin{aligned}
\mathrm{SIC}_{\mathrm{error}}(t) &= \Delta^2_{A,B}P(T>t|\mathrm{ER}) \\
&= \Delta^2_{A,B}\left[P(T>t\cap\mathrm{ER})\times\frac{1}{P(\mathrm{ER})}\right] \\
&= \Delta^2_{A,B}\left[P(T>t\cap[(\neg A,\neg B)\oplus(\neg A,B)\oplus(A,\neg B)])\times\frac{1}{P(\mathrm{ER})}\right] \\
&= \Delta^2_{A,B}[P(T>t\cap(\neg A,\neg B))+P(T>t\cap(\neg A,B)) \\
&\quad +P(T>t\cap(A,\neg B))]\times\frac{1}{P(\mathrm{ER})} \\
&= \Delta^2_{A,B}[S(t|\neg A,\neg B)\times(1-p_A)(1-p_B) \\
&\quad +S(t|\neg A,B)\times(1-p_A)p_B \\
&\quad +S(t|A,\neg B)\times p_A(1-p_B)]\times\frac{1}{P(\mathrm{ER})}.
\end{aligned}$$

Thus, we have the same problem as for the correct OR task results in that we cannot presume the probability that the system is in any of these error states for either the Parallel Exhaustive or Parallel Self-Terminating model and rely on simulations to investigate their shape.

Serial Models

The final three states in Equations 8 and 9 generate error responses. The error conditioned SIC reflects a mixture of error states preventing a proof of the SIC shape. We again rely on simulations, presented below, to investigate these models.

OR Task

In an OR task, error responses are generated whenever both channels fail to detect the presence of a target. As we shall see, errors in an OR task require exhaustive responding, consequently allowing proofs of the SIC shapes conditioned on error responding.

Parallel Models

The final state in Equations 5 and 14 is the sole error state in the Parallel Exhaustive and Parallel Self-Terminating models, respectively, in the OR task. If we assume that the distributions of the error RTs are stochastically ordered such that $S_L(t|ER)$ dominates $S_H(t|ER)$, then we can proceed with a proof of the error conditioned SIC shape.

Proposition 5. Assuming independence, selective influence, and stochastic dominance of the error RTs, the Parallel Exhaustive model, in an OR task, predicts that the conditional error SIC is negative for all t.

Proof. The proof follows the same form as for the conditional correct SIC for the parallel models in the AND task but with the error accumulator taking the place of the correct accumulator. □

Proposition 6. Assuming independence and selective influence, along with stochastic dominance of the error RTs, the Parallel Self-Terminating model, in an OR task, predicts that the conditional error SIC is negative for all t.

Proof. The proof follows the same form as for the conditional correct SIC for the parallel models in the AND task and the conditional error SIC in the OR task. □

Serial Models

The error state in both Serial models arises through summing the finishing times from both accumulators when both accumulators fail to detect both targets. Hence, if we assume stochastic dominance of the error RTs, then the proof of the SIC shape can proceed.

Proposition 7. Assuming independence, selective influence, and stochastic dominance of the error RTs, the Serial Exhaustive model in an OR task predicts that the conditional SIC for error responses is negative for small values of t and later becomes positive. Furthermore, the integral over the positive real line (i.e., of t) of the SIC function is 0.

Proof. The error conditioned SIC for the Serial Exhaustive model in an OR task follows the same form as the correct conditioned SIC for the Serial models in the AND task with the error accumulator replacing the correct accumulator. □

Proposition 8. Assuming independence, selective influence, and stochastic dominance of the error RTs, the Serial Self-Terminating model, in an OR task, predicts that the conditional SIC for error responses is negative for small values of t and later becomes positive for larger values of t. The integral over the positive real line (i.e., of t) of the SIC function is 0.

Proof. The proof is the same as for the Serial Exhaustive model with the error accumulator replacing the correct accumulator. □

Interim Summary II

For models in which the response is generated from a single state, we have provided proofs that, conditioned on correct and error responding, the SIC functions take on the same qualitative shape as their error-free counterparts. For the models in which the response is generated from a mixture of states, however, because we are unable to identify the probability that the system is in each of the correct states, we are unable to determine the shape of the conditional SIC in the usual general fashion.

We next explore the relationship between high and low salience further by analyzing the behavior of correct and error RTs resultant from a salience manipulation applied to the rates of two major classes of parameterized process models: the Wiener diffusion sequential sampling model and its current extensions and Poisson counter-race models.[1] Although the latter do have some popular instantiations, they have not been so standardized as have the usual diffusion models. It will be seen that even the more complex diffusion models, with mixtures of starting states and drift rates deliver canonical types of dominance relations and thereby the common SIC predictions.

Intriguingly, the Poisson race models, depending on exact assumptions about the rates of correct versus incorrect processing channels, can result in distinct dominance relationships on "wrong" decisional RTs. In fact, we prove that when the rates are constrained to equal the same sum across different levels of salience (i.e., the sum of the rates for the correct and incorrect accumulators are the same for a high salience stimulus as it is for a low salience stimulus), the correct RTs will be faster for the high salience case compared to the low salience case. By contrast, the error RTs will be faster for the low salience case compared to the high salience case. When the salience manipulation only affects the correct channel, we show that high salience correct RTs are still faster than low salience correct RTs, and that high salience error RTs are also faster than low salience error RTs. This will also be the prediction for Poisson race models where both rates can change with salience but the difference between correct rates will be larger than than between incorrect rates.

Selective Influence and Stochastic Dominance for Specific Popular Parameterized Models

The original SIC proofs (Townsend & Nozawa, 1995) and the new, conditional SIC expressions above rely on the critical assumption of effective selective influence, namely that the speed of processing a high salience stimulus (H) is faster than the speed of processing a low salience stimulus (L; i.e., that the manipulation of salience is effective) and that the manipulation of Channel A had no impact on Channel B (i.e., that there is selective influence of each channel). This leads to a dominance relationship on the RT distributions (Townsend, 1990b), as discussed earlier. To extend the SIC to systems that may give rise to errors, this salience assumption needs to be discussed for correct and errors separately.

Here we explore the relationship between high and low salience by analyzing the behavior of correct and error RTs resultant from a salience manipulation applied to the rates of the two major classes of process models mentioned above, the Wiener diffusion sequential sampling models and extensions and two versions of Poisson counter-race models. We will demonstrate that even the fairly complex diffusion models make the canonical qualitative predictions engendered in the error-free case for both the dominance relations and the SIC predictions.

Moving to Poisson counter-race models, three versions are taken up, that differ in the ways processing rate changes across correct and incorrect accumulators: (a) Poisson Model I: It is assumed that the sum of the correct and incorrect processing rates stays constant as salience is manipulated (such that a higher-quality signal leads to faster correct rate and slower incorrect rate). (b) Poisson Model II: It is assumed that while the correct rate is increased as a function of salience, the incorrect rate stays constant. (c) Poisson Model III: Both incorrect and correct rates can change and the difference between the correct and incorrect rates becomes larger with increased salience but the sum is unconstrained.

Next, intriguingly, we will learn that the more frequently utilized Poisson race model, Model I, does make the expected prediction that processing times conditioned on being correct are ordered such that correct decisions are faster when salience is higher. In surprising contrast, this class of models predicts that incorrect decisions are actually faster in a strong stochastic sense, when salience is lower! The opposite is the case for Poisson race Models II and III: both correct and incorrect decisions are faster with higher salience.

Diffusion models assume a kind of random walk back and forth on continuous evidence regarding the two choices (Ratcliff, 1978). Although discrete state-space random walks are possible (e.g., Link & Heath, 1975) and random walk models based on Poisson distributions have now appeared (Blurton et al., 2020), we focus on the continuous state-space models here with the usual underlying Wiener process noise system.

On the other hand, in race models of which Poisson models are the most common, two channels, each comprised of two counters, are separately accumulating evidence for the two respective alternatives. The most studied of these assume independent Poisson counting distributions in the two channels (e.g., Audley & Pike, 1965; Townsend & Ashby, 1983). However, generalizations to non-Poisson counters have been increasing in frequency (e.g., Smith & Van Zandt, 2000; Tillman et al., 2020; Townsend & Liu, 2020).

All these models can be characterized as sequential sampling models and all are useful because of their ability to provide predictions for correct and error RTs, and because in many situations, these models provide a good explanation of various data patterns using distinct parameters. Examples include but are not limited to speed–accuracy trade-off, bias toward one choice, the skewness of RT data, and the relative speeds of correct and error responses (Luce, 1986; Ratcliff et al., 1999; Ratcliff & Rouder, 2000; Van Zandt & Ratcliff, 1995). Computational methods have been developed to obtain predictions for such models (Diederich &

[1] Readers should take note that the diffusion and Poisson models are being applied to individual channels. Thus, a diffusion model of the full response would involve two diffusion processes—one for each channel—arrayed in serial or in parallel.

Busemeyer, 2003; Smith & Van Zandt, 2000). We provide predictions deriving directly from the analytic forms of the RT expressions, rather than using numerical methods.

Sequential sampling models assume that decisions are made by accumulating noisy information to some predetermined decision criteria. Therefore, performance in a task is affected by two main factors: the quality of the information extracted from the stimulus which determines the accumulation rate (i.e., the stimulus salience level) and the quantity of information required to make a response. The value of these two factors combined together produces predictions on RTs and correctness of each trial. In this article, we limit our scope to the effect of salience manipulation.

We consider conditional RTs (conditioned on their correctness) as random variables that will change as a function of the salience level of the stimulus. To examine the effect of salience, H or L, on the conditional distributions, we will rely on hierarchical inference properties outlined by Townsend (1990b). We briefly reviewed some major aspects of that hierarchy in an earlier section.

In the following, we discuss the effect of salience manipulation in terms of RTs based on the two classes of sequential sampling models. Specifically, we discuss how correct and incorrect RTs change under different salience conditions, and how strong the effect is in terms of the hierarchical inference theory. Rather remarkably, it turns out felicitously that dominance will hold at the strongest level in the hierarchy and we take advantage of that fact in our proofs. We start with diffusion model.

Diffusion Model

To motivate our article, we have used the example of a pair of accrual halting models (see Figure 4); however, our proofs rely only on the conditional distributions of each channel and not on the completion time distributions of individual accumulators. Consequently, we now investigate diffusion models which are not independent race processes within each channel. Instead, diffusion models assume that the relative evidence for each outcome is accumulated continuously with each increment a draw from a Gaussian distribution. A sample path from a diffusion model is depicted in Figure 6.

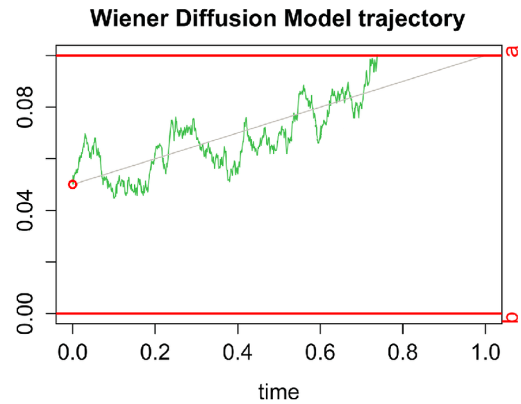
The *first passage time*, that is the time for a drifting particle to hit either boundary a or b, is given by the following equation for constant a and b:

$$f(t|R_a) = \frac{1}{P(a)} \frac{\pi s^2}{(a-b)^2} \exp\left[\frac{v(a-z)}{s^2} - \frac{v^2 t}{2s^2}\right] \times \sum_{k=1}^{\infty} k \exp\left[-\frac{k^2\pi^2 s^2 t}{2(a-b)^2}\right] \sin\left[\frac{k\pi(a-z)}{(a-b)}\right]. \quad (17)$$

where v is the drift rate, s is the drift coefficient (s^2 is commonly termed the diffusion rate), and z is the starting point of accumulation (cf. Ratcliff, 1978). The boundaries a and b represent different response options in a two-alternative forced-choice (2AFC) task. In the case, where termination at the a boundary is correct, one can evaluate the first passage time density for error responses terminating at b by replacing, $a - z$ with $z - b$, and $P(a)$ with $P(b)$. In the case where termination at b is correct, $f(t|R_b)$ can be found by setting the sign of v to be negative. In this latter case, the first passage time for error response is found by evaluating $f(t|R_a)$.

Figure 6

Illustration of the Wiener Diffusion Model With No Bias



Note. Response boundaries are labeled a and b. $v = 0.05$ is the drift rate. See the online article for the color version of this figure.

$P(a)$ and $P(b)$ give the probability of terminating at the upper, a, or lower, b, boundaries, respectively. $P(a)$ is computed as:

$$P(a) = \frac{\exp(-2vz/s^2) - \exp(-2vb/s^2)}{\exp(-2va/s^2) - \exp(-2vb/s^2)}. \quad (18)$$

In order to account for empirical data, in particular patterns of correct and error RTs, several parameters are assumed to vary across trials (Laming, 1968; Ratcliff, 1978; Ratcliff & Rouder, 1998). Here we consider two common sources of variability. The first is across-trial variability in the starting point of accumulation, z. The second is across-trial variability in the drift rate, v, which is assumed to be normally distributed across trials with mean, μ_v, and standard deviation, σ. Such models have been termed the DDM (Ratcliff & McKoon, 2008).

In the following, we will present several results on the dominance relations between a high salience condition model, H, and a low salience condition model, L, by varying the drift rate from higher to lower, respectively. We will start from the perspective of the Wiener diffusion model and then extend to models with different types of across-trial variability. For each pair of item conditions, H and L, we will present predictions for correct and error RTs separately.

In Appendix A, we prove that the RT distribution conditioned on correct responding for the high salience target dominates the low salience target for all t. The proof follows from the properties of the likelihood function and the relationship between the likelihood functions and the CDFs for the H and the L cases (see Proposition 9). We can prove an additional result that $F_H(t|R_b) > F_L(t|R_b)$ using the same strategy. The implication of this result is that both correct and incorrect responses are slower for the low salience stimulus, L, than for the high salience stimulus, H. That is, the required dominance relationship holds.

The consistency between correct and error responses is not surprising. In fact, in the absence of bias (i.e., when $a - z = z - b$), Wiener diffusion models predict that the conditional RT distributions for correct and incorrect responses will be identical. This property renders the Wiener diffusion model unsuitable as a model of human decision-making, where it is common to observe error RTs which are faster or slower than correct RTs.

To remedy this issue, across-trial variability in the starting point and drift rate are added to the model. It is common to assume that starting point is a uniform random variable, $z \sim U(z_0 - h, z_0 + h)$, where 2 h is the width of the Uniform distribution. Drift rate is commonly assumed to be a Gaussian random variable, $v \sim N(\mu_z, \sigma_z^2)$. Under these assumptions, the predicted first passage time distribution is a compound distribution that results from compounding the Wiener diffusion model with the parameter distribution.

In Appendix A, we show generally that with each type of RT variability, the compound distribution for the high salience target, $\bar{F}_H(t)$, dominates the compound distribution for the low salience target, $\bar{F}_L(t)$ (see Propositions ?? and 10). These proofs suggest that, for Wiener diffusion models, stochastic dominance of H over L is unaffected by across-trial variability in either starting point or drift rate. Proposition 11 shows that this result holds for the general model in which there is both starting point and drift rate variability.

These proofs demonstrate that the diffusion model provides consistently ordered predictions with regards to drift rates representing high and low salience conditions substantiating selective influence. Even with variability in the drift rate and variability in the starting point, the diffusion model can only predict one ordering: RT increased with lower salience drift rate, and this prediction holds for both correct and incorrect responses. Thus, this entire class of models will produce the canonical SIC patterns under the appropriate stopping rules.

Poisson Race Models

The basic Poisson race model (A. R. Pike, 1966; LaBerge, 1994; Logan, 1996; R. Pike, 1973; Smith & Van Zandt, 2000; Townsend & Ashby, 1983; Van Zandt et al., 2000), also called the Poisson counter model, was propagated by the assumption that neural spikes are emitted when a stimulus is presented. These spikes are independently stored in different internal "counters" representing each alternative choice. A response is triggered when the number of spikes reaches a threshold level for either alternative. An important pair of parameters in the Poisson counter model is the time between arrival of two spikes in each accumulator. In the Poisson race model, the time between the arrival of two spikes follows an exponential distribution with rate, v. This assumption implies that the accumulation process in each counter is a Poisson process. This type of model remains of interest regardless of whether it properly depicts neural spike events.

We label the two counters as correct, C, and incorrect, I, and we label the two rates as v_C and v_I. Subject bias is reflected in the criteria value, K, of each counter. Increased bias to one alternative will be reflected by a smaller value of K for that alternative. Here, we assume that there is no subject bias. Townsend and Ashby (1983; Chapter 9) explored model predictions for processing time statistics of correct and incorrect responses. For instance, in several notable Poisson counting models, corrects were faster than errors. Smith and Van Zandt (2000) demonstrated that inhomogeneous Poisson counters with proportional counting rates always predict faster mean processing times for correct than for incorrect, decisions. Townsend and Liu (2020) have very recently proven that, in general, whether corrects are faster than errors in independent race models (not confined to, e.g., Poisson counters), depends on the underlying hazard functions. Futhermore, they extended the conclusions from Smith and Van Zandt (2000) on the mean processing times for races with proportional hazard functions, but otherwise inhomogeneous Poisson counters, to conditional distribution function orderings. Finally, they discovered the existence of inhomogeneous Poisson counters that could produce faster incorrect than correct responses.

In the present investigation, we stick to ordinary Poisson models. We will use $g(t)$ and $G(t)$ for the density and distribution functions. The density functions for the RT are given below. Without loss of generality, we assume the correct counter is associated with response a, thus $g(t|R_a)$ is the density for correct responses, and $g(t|R_b)$ is the density for the error responses.

$$g(t|R_a) = \frac{1}{P(a)} \times \frac{(v_C t)^{K-1} v_C \exp(-v_C t)}{(K-1)!} \times \sum_{j=0}^{K-1} \frac{(v_I t)^j \exp(-v_I t)}{j!} \tag{19}$$

$$g(t|R_b) = \frac{1}{P(b)} \times \frac{(v_I t)^{K-1} v_I \exp(-v_I t)}{(K-1)!} \times \sum_{j=0}^{K-1} \frac{(v_C t)^j \exp(-v_C t)}{j!}, \tag{20}$$

in which $P(a)$ and $P(b)$ are the response probabilities.

The rates of the two Poisson processes determine the response probability and the RT. With regard to our salience manipulation, a mild assumption is the translation of the physical stimulus as a nondecreasing function of the evidence for its corresponding response; that is, that the rate in the correct counter is higher in the high salience condition than in the low salience condition; hence, $v_C^H > v_C^L$.

As with the diffusion model, we assume that the salience manipulation affects only the processing rates and not the other parameters.[2] A second important factor that affects both response probability and RT is the relative relationship between the rates of the two counters. Here we divide the set of possible relationships into three cases and discuss the RT predictions for correct and error responses for each case separately. The three cases are as follows:

1. Poisson Race Model I: When the sum of the rates in both counters is constant over the salience manipulation, $v_C^H + v_I^H = v_C^L + v_I^L$. That is, when there is a trade-off between the correct and error accumulator rates. Such a trade-off would occur, for instance, if the rates were constrained to sum to one across both accumulators, which is a common assumption in models of RT (Fific et al., 2010; Little et al., 2016; Nosofsky & Palmeri, 1997b).

2. Poisson Race Model II: When the salience manipulation only affects the rate in the correct counter, $v_C^H > v_C^L$ but the rates are equal for the error counter, $v_I^H = v_I^L$. Such a condition might arise in a detection task, where positive evidence for the presence of a target would be naturally larger at higher salience than at lower salience, but the

[2] We do not consider across-trial variability in the Poisson model as the results of the Poisson model can be distinguished from the diffusion model without the additional variability. Note, however, that Proposition 3 is quite general and applies also to the Poisson model with across-trial variability.

absence of evidence would be equated across both high and low equations.

3. Poisson Race Model III: When the salience manipulation affects the change of the rates in each counter, $v_C^H - v_C^L > v_I^H - v_I^L$. That is, when the rates are greater for the high salience conditions in both the correct and error counter and the difference is greater in the correct accumulator than in the incorrect accumulator. This model assumes that the change from L to H increases the rates in both the correct and error accumulators. This is a somewhat unusual model in that one would usually expect an increase in salience to decrease the rate for the error accumulator, as discussed earlier. Nevertheless, we retain the case here to examine the dominance behavior of this assumption. As defined by this inequality, Model III includes Models I and II as special cases.

Under the first case, Model I, which is the most intuitive of the three, we show in Appendix B that the Poisson counter model predicts faster correct RTs in the high salience condition but faster error responses in the low salience condition (see Propositions 12 and 13). This is the opposite of the prediction made by the diffusion model. The intuition is that the correct rate is larger in the high salience condition than in the low salience condition; combined with the fact that the sum of the rates has to be equal, the implication is that the error rate has to be higher for the low salience condition compared to the high salience condition. Consequently, we prove that while the dominance of H over L holds for correct responses, [i.e., $G_H(t|Correct) > G_L(t|Correct)$], it is reversed for error responses, [i.e., $G_H(t|Incorrect) < G_L(t|Incorrect)$; where $G(t)$ is the CDF of the Poisson counter model; see Appendix B]. Figure 7 presents simulated mean RT and accuracy data comparing Model I Poisson counter to the diffusion model with across-trial variability. The implication of this result is that for error responses, under the assumptions listed in Model I, the SIC shapes may *not* hold to their error-free counterparts.

In Race Models II and III, we show that the Poisson counter model predicts faster RTs in the high salience condition than in the low salience conditions for both the correct and error responses (see Propositions 14–17). Under these conditions, the SIC results *will hold* for both correct and error RTs.

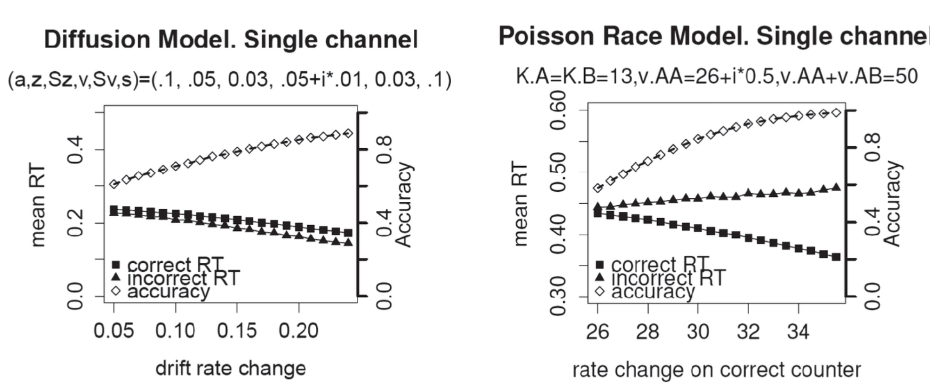
Interim Summary III

Our theorems for the two basic types of process models capable of making predictions for both RTs and accuracy show that, while the dominance assumption for correct RTs holds for both the Wiener diffusion and Poisson counter models, this is not the case for error RTs. Although low salience stimulus error RTs are slower than high salience stimulus error RTs for the diffusion model, this is only the case for some parameterizations of the Poisson counter model, namely Race Models II and III. In the final section of our article, we will provide demonstrations of the conditional SIC shapes using distributions derived from the Poisson counter model.

Computation of Conditional SIC Functions for Specific Distributions

To demonstrate that the shapes of the conditional SIC functions follow the form of the error-free SICs when the stochastic dominance assumption is met, we simulated a Poisson race model using Equations 5, 8, 14, and 16 to compute the RTs for correct and error trials. We conducted two simulations: one using Poisson Race Model I in which the sum of the correct and incorrect accumulator rates was constant across the salience manipulation (Simulation 1) and one using Poisson Race Model II in which the salience manipulation only affected the rate for the correct accumulator (Simulation 2). In the first simulation, we set the total of the A accumulators to 50 and the total of the B accumulators to 44 to capture the idea that each pair might not have the same distribution. Likewise, for Simulation 2, the error accumulator total was set to a constant 15 for pair A and 12 for pair B. We used four levels of salience manipulation; the rate parameter values are shown in Table 2.

Figure 7
XXX



Note. Left: Simulated mean RTs from a Wiener diffusion model with across-trial variability. a = upper boundary, z = starting point, Sz = measure of across-trial variability on z (half of the length on the uniform distribution's support), v = drift rate, Sv = measure of across-trial variability on v (standard deviation of drift rate), s = scaling parameter. Right: Simulated mean RTs from Poisson Model I. K.A, K.B = response criteria, v.AA = processing rate for the correct counter, v.AB = processing rate for the incorrect counter

Table 2
Rates for Poisson Accumulator Simulation of Parallel and Serial Exhaustive Models in an AND Task

	High salience				Low salience			
	AC	AI	BC	BI	AC	AI	BC	BI
	Simulation 1 Sum held constant							
Level 1	45	5	42	2	40	10	37	7
Level 2	40	10	37	7	35	15	32	12
Level 3	35	15	32	12	30	20	27	17
Level 4	30	20	27	17	25	25	22	22
	Simulation 2 Salience only affects correct							
Level 1	45	15	42	12	40	15	37	12
Level 2	40	15	37	12	35	15	32	12
Level 3	35	15	32	12	30	15	27	12
Level 4	30	15	27	12	25	15	22	12

Across both simulation sets and all four levels, the criteria parameters were set to 13 across all accumulators. We simulated all combinations of the Exhaustive and Self-Terminating Parallel and Serial models in the AND task and the OR task. Note that for correct responses in the AND task, Exhaustive and Self-Terminating models make identical predictions. The SIC results are conditioned on correct responding. Poisson model simulation code is available at: https://github.com/knowlabUnimelb/SFTERRORS.

Parallel and Serial Exhaustive Models in an AND Task

The accuracy rates from each simulation for the Parallel Exhaustive model are shown in Table 3 for each stimulus: HH, HL, LH, and LL. As shown in Table 3, Simulation 1 resulted in very low accuracy, particularly for the LL stimulus (as low as 25%), while Simulation 2 resulted in modest error rates for all of the stimuli including the LL stimulus (accuracy of 84% or better). As shown in Figure 8, the conditional SIC function is completely negative for the Parallel Exhaustive model—consistent with the predictions of the error-free SIC function. Visual inspection of the conditional survivor functions shows that stochastic dominance is preserved for the double-target items.

The mean RTs and accuracy rates for the Serial Exhaustive model are shown in Table 4 and the conditional SIC functions from the simulated Serial Exhaustive model are shown in Figure 9. For the Serial Exhaustive model, the simulated accuracy rates are comparable to the Parallel Exhaustive model, but the overall mean RTs are longer for the Serial model as might be expected. As shown in Figure 9, the accuracy-conditioned SICs show the "S" shape characteristic of the error-free SIC. Regardless of the accuracy rate, each function was negative for early times and positive for later times with the integral of the function equalling zero. Stochastic dominance is again preserved for the double-target items (see top panels of Figure 9).

Parallel and Serial Models for OR Designs

Parallel Self-Terminating Model in an OR Task

We next investigated the shape of the Parallel Self-Terminating conditional SIC using simulations. The simulated rate values are reported in Table 5. Visual inspection of the survivor functions in Figure 10 reveals that stochastic dominance holds in all cases of Simulation 2 and most cases, save Level 4, for Simulation 1. In Level 4 of Simulation 1, the survivor functions start to converge and although stochastic dominance still holds, there is more overlap than in the other cases (Table 6).

The simulated SICs are shown in Figure 10. For Simulation 1, at very low levels of accuracy, the SIC is near 0 demonstrating the fact that the SIC is a mixture of different states with mixture probabilities depending on different error levels. It is possible, though we do not see it here, that at even lower accuracy, for Simulation 1, the survivor functions cross over in ways which render the conditional SIC uninterpretable. For Simulation 2, where the accuracy is higher across salience levels, the SICs are always positive for all t. Furthermore, the maximum deviation from zero increases with decreasing accuracy. Overall, the results of both simulations are consistent with the error-free SIC results, showing a completely positive conditional SIC function.

Serial Self-Terminating Models in an OR Task

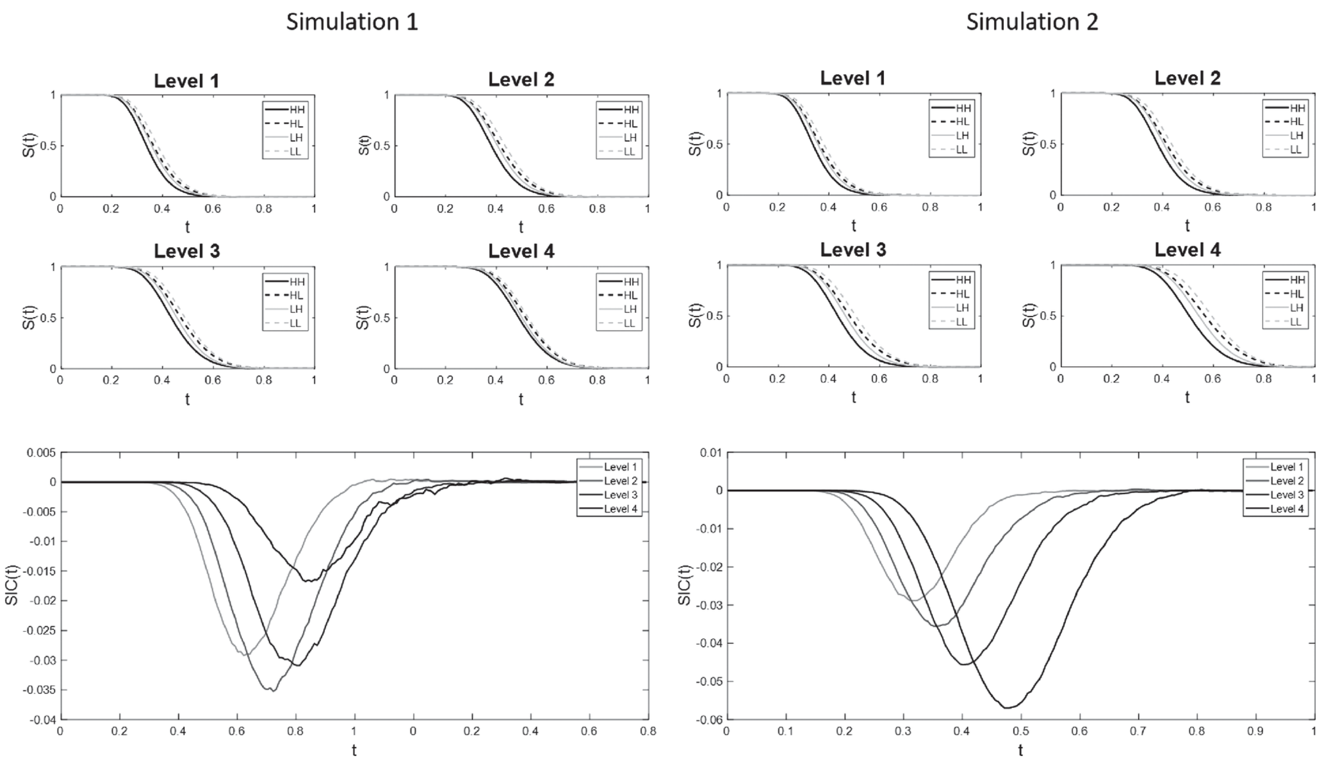
We simulated the Serial Self-Terminating model using the parameters shown in Table 5. The mean RT and accuracy results are shown in Table 7. For both simulations, there exists some conditions under which the conditional SIC does not match the expected form (see

Table 3
Parallel Exhaustive Mean RTs and Error Rates in an AND Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	346	1.00	392	1.00	449	0.97	501	0.74
HL	376	1.00	430	1.00	488	0.86	527	0.42
LH	366	1.00	416	0.98	469	0.85	513	0.44
LL	391	1.00	448	0.97	502	0.75	537	0.25
				Simulation 2				
HH	347	1.00	391	0.99	447	0.97	522	0.94
HL	376	0.99	429	0.98	501	0.96	597	0.89
LH	366	0.99	416	0.98	478	0.95	563	0.88
LL	391	0.99	448	0.97	523	0.94	622	0.84

Note. RT = response time.

Figure 8
Parallel Exhaustive AND Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

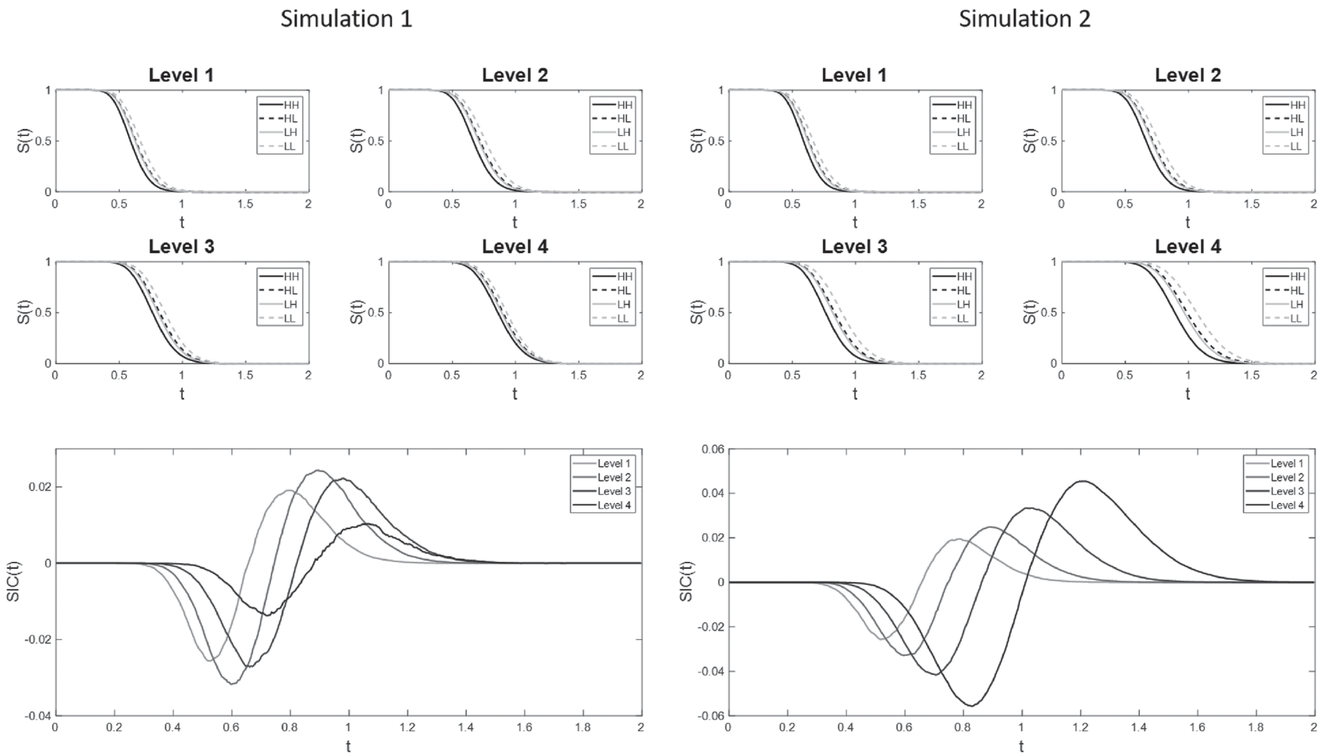
Figure 11). While the SIC simulations with near ceiling accuracy (e.g., Level 1) are approximately zero for all *t*, as the accuracy rate decreases the function begins to deviate from zero. In some cases, the SIC shows the expected negative to positive pattern, but inspection of the survivor functions reveals a violation of stochastic dominance in the cases where the SIC deviates from the expected error-free form. The implication is that, once conditioned on accuracy, there is a clear violation of stochastic dominance and therefore selective influence, which renders the SIC uninterpretable. Hence, the stochastic dominance of the double-target conditional survivor functions (LL > LH, HL > HH) needs to be assessed very carefully before interpreting the conditional SIC. This can be assessed by plotting the survivor functions and examining these plots for crossovers; however, more sophisticated statistical tests are available to examine stochastic dominance (Houpt et al., 2014). In the present simulations, the conditional SIC functions for the Serial Self-Terminating case are not interpretable, in line with the clear violations of stochastic dominance.

Table 4
Serial Exhaustive Mean RTs and Error Rates in an AND Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	598	1.00	678	1.00	773	0.97	870	0.74
HL	641	1.00	730	0.99	827	0.86	909	0.43
LH	635	1.00	721	0.98	814	0.84	897	0.43
LL	678	1.00	773	0.97	869	0.74	936	0.25
				Simulation 2				
HH	597	1.00	674	0.99	772	0.97	904	0.94
HL	640	0.99	728	0.98	844	0.96	1,003	0.90
LH	632	0.99	720	0.98	830	0.95	979	0.88
LL	674	0.99	773	0.97	903	0.94	1,077	0.84

Note. RT = response time.

Figure 9
Serial Exhaustive AND Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

Parallel Exhaustive Model in an OR Task

For our final two simulations, we Simulated Exhaustive model performance in an OR task using Equations 13 and 15. Correct responding in these models is also the result of a mixture of states; however, the predicted RT within each state is an exhaustive process of both channels. We used the same simulated Poisson accumulator rates as for the Self-Terminating models in the OR task (see Table 5).

Table 5
Rates for Poisson Accumulator Simulation of Parallel and Serial Models in an OR Task

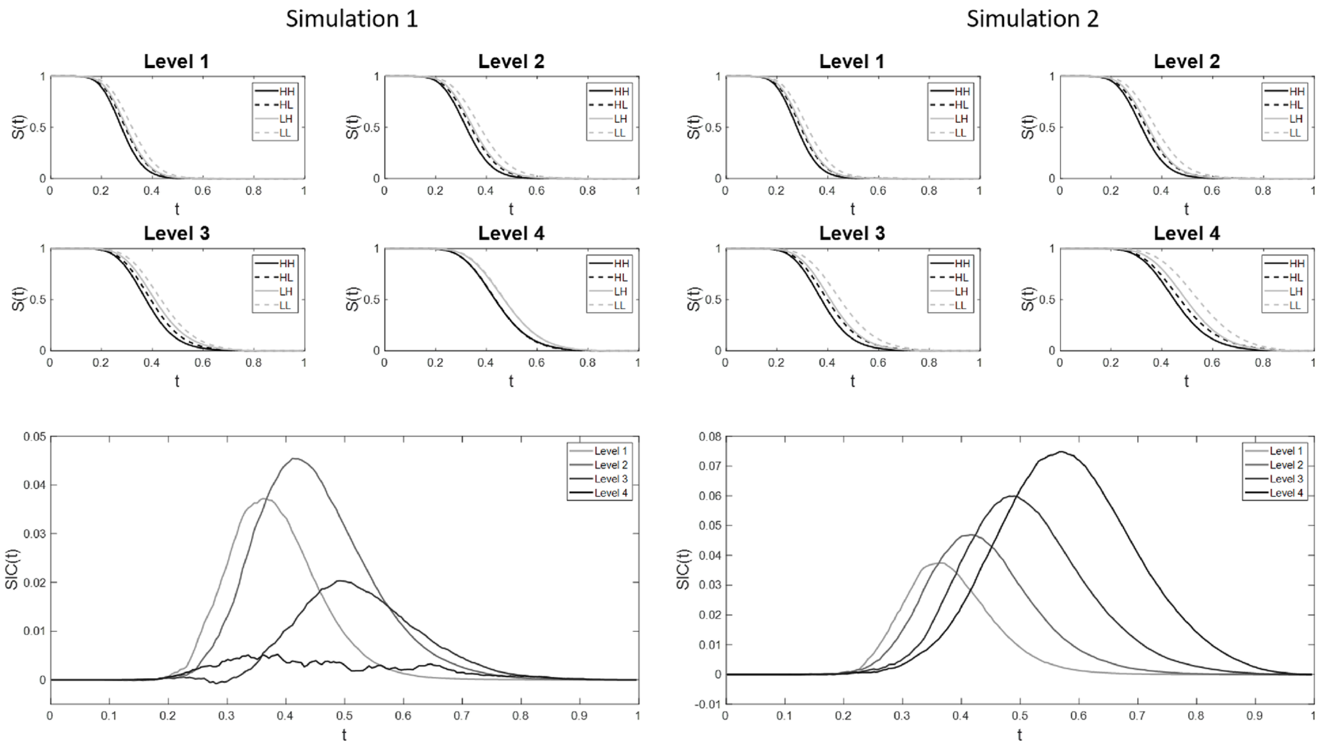
	High salience				Low salience			
	AC	AI	BC	BI	AC	AI	BC	BI
	Simulation 1 Sum held constant							
Level 1	40	10	37	7	35	15	32	12
Level 2	35	15	32	12	30	20	27	17
Level 3	30	20	27	17	25	25	22	22
Level 4	25	25	22	22	20	30	27	20
	Simulation 2 Salience only affects correct							
Level 1	40	20	37	17	35	20	32	17
Level 2	35	20	32	17	30	20	27	17
Level 3	30	20	27	17	25	20	22	17
Level 4	25	20	22	17	20	20	17	17

The accuracy rates from each simulation for the Parallel Exhaustive model are shown in Table 8 for each stimulus: HH, HL, LH, and LL. The accuracy rates were comparable to the Parallel Self-Terminating model, but the simulated RTs were larger as expected. The survivor functions in Figure 12 are well ordered with the exception of Level 4 in Simulation 1. Here, the survivor function of the LL double target overlaps the HL double target. The simulated SICs are shown at the bottom of Figure 12. For all of the simulations, the SIC is negative for all t. That is, despite the overlap with lower levels of accuracy in Level 4 (Simulation 1), the accuracy-conditioned SIC matches is the error-free SIC. This is likely due to the fact that the maximum time is still much slower for the HL, LH, and LL double targets compared to the HH double target even when mixing different accuracy states. For the remaining simulations, which preserve the stochastic dominance of the double targets, the conditional SIC results are completely negative like their error-free counterpart.

Serial Exhaustive Model in an OR Task

Finally, for the Serial Exhaustive model, we again found substantial crossover in the survivor functions of the double-target stimuli (see Figure 13). The mean RT and accuracy results are

Figure 10
Parallel Self-Terminating OR Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

shown in Table 9. The stochastic dominance was violated most severely in Simulation 1, where the LH stimulus was often slower than the LL stimulus. One in the Level 1 simulation, where accuracy was at ceiling, was the order of the double targets preserved. In this case, the accuracy-conditioned SIC looks like its error-free counterpart (as shown in Figure 3). For Simulation 2, based on visual inspection of Figure 13, stochastic dominance is mostly preserved. The conditional SICs have the characteristic S-shape (initially negative and then positive) though the size of the positive region becomes smaller as accuracy decreases. This can also be confirmed by using the mean RTs in Table 9 to compute the MIC using Equation 1. As accuracy decreases, the MIC becomes more negative. At lower levels of error in an OR task, Serial Exhaustive processing may mimic Parallel Exhaustive processing at the level of the conditional SIC. Though, where the Parallel Exhaustive conditional SIC is completely negative, the Serial Exhaustive conditional SIC shows some positivity at later times.

Table 6
Parallel Self-Terminating Mean RTs and Error Rates in an OR Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	285	1.00	327	1.00	386	.98	446	.75
HL	300	1.00	347	1.00	405	.92	446	.57
LH	305	1.00	355	1.00	425	.94	480	.58
LL	327	1.00	386	.98	447	.75	481	.26
				Simulation 2				
HH	285	1.00	328	1.00	386	.98	463	.93
HL	301	1.00	349	.99	412	.92	492	.86
LH	305	1.00	355	.99	421	.97	511	.87
LL	328	1.00	386	.98	464	.93	567	.75

Note. RT = response time.

Table 7
Serial Self-Terminating Mean RTs and Error Rates in an OR Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	326	1.00	379	1.00	494	.98	661	.75
HL	326	1.00	379	1.00	465	.92	514	.57
LH	378	1.00	495	1.00	730	.94	990	.58
LL	379	1.00	495	.98	660	.75	760	.26
				Simulation 2				
HH	339	1.00	401	1.00	494	.98	635	.93
HL	341	1.00	403	.99	493	.96	608	.86
LH	399	1.00	491	.99	636	.97	866	.87
LL	402	1.00	495	.98	635	.93	933	.75

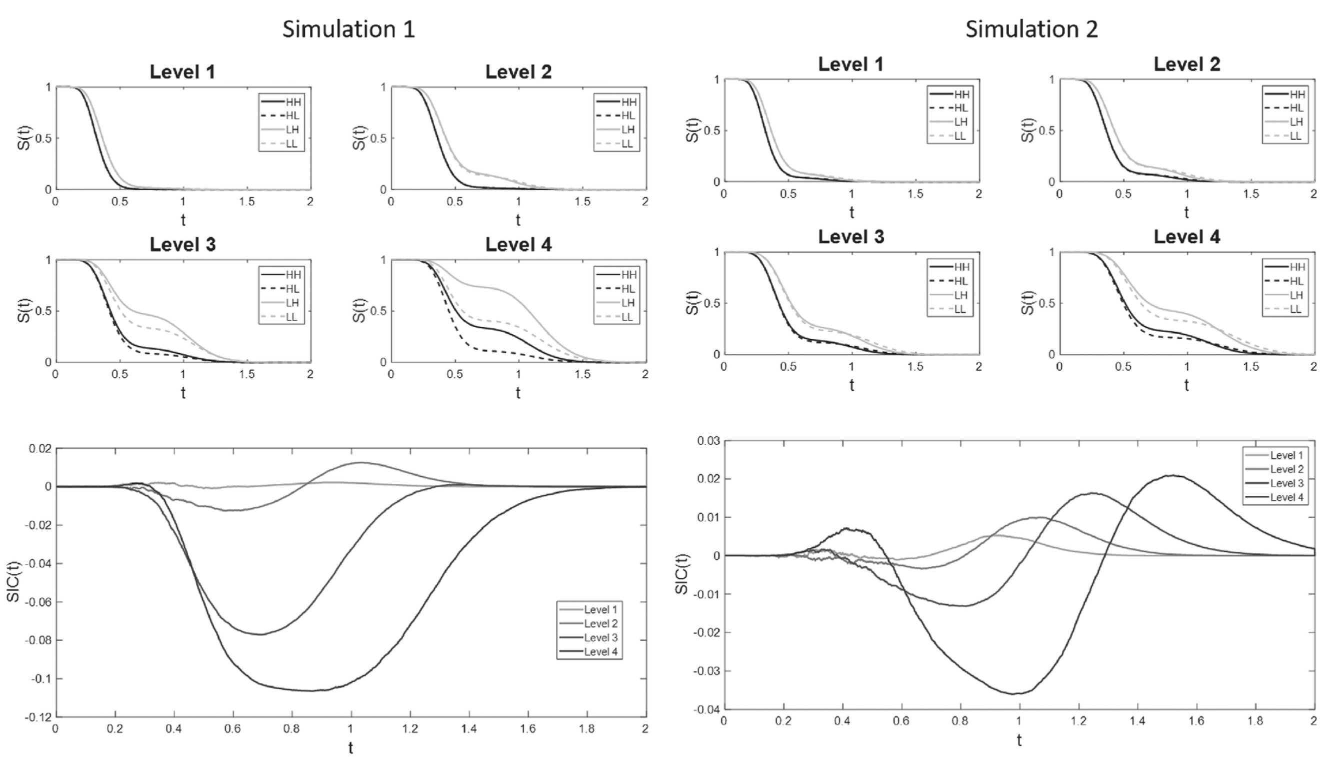
Note. RT = response time.

General Discussion

This article extends the SFT methodology to account for RTs conditioned on correct and error responding. A critical assumption of these results is that the double-target stimulus conditions are ordered in a manner reflecting the effective selective influence of the experimental manipulation. With satisfaction of this assumption, we have shown that the SICs for these conditional RTs, under exhaustive processing, maintain their diagnosticity regardless of the level of error. We showed through formal proofs that a diffusion model with across-trial drift and starting point variability preserves the dominance relationship when it is assumed that a high salience signal has a higher drift rate than a low salience signal.

We likewise proved that the stochastic dominance relationship holds for two classes of Poisson counter model. One in which increasing salience increases the rate of the correct accumulator only

Figure 11
Serial Self-Terminating OR Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

Table 8
Parallel Exhaustive Mean RTs and Error Rates in an OR Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	392	1.00	451	1.00	529	.98	619	.75
HL	432	1.00	507	1.00	607	.92	757	.56
LH	419	1.00	486	1.00	573	.94	681	.58
LL	451	1.00	529	.98	620	.75	747	.26
				Simulation 2				
HH	392	1.00	450	1.00	528	.98	636	.93
HL	431	1.00	506	.99	610	.96	767	.86
LH	418	1.00	485	.99	575	.97	703	.87
LL	451	1.00	529	.98	636	.93	792	.75

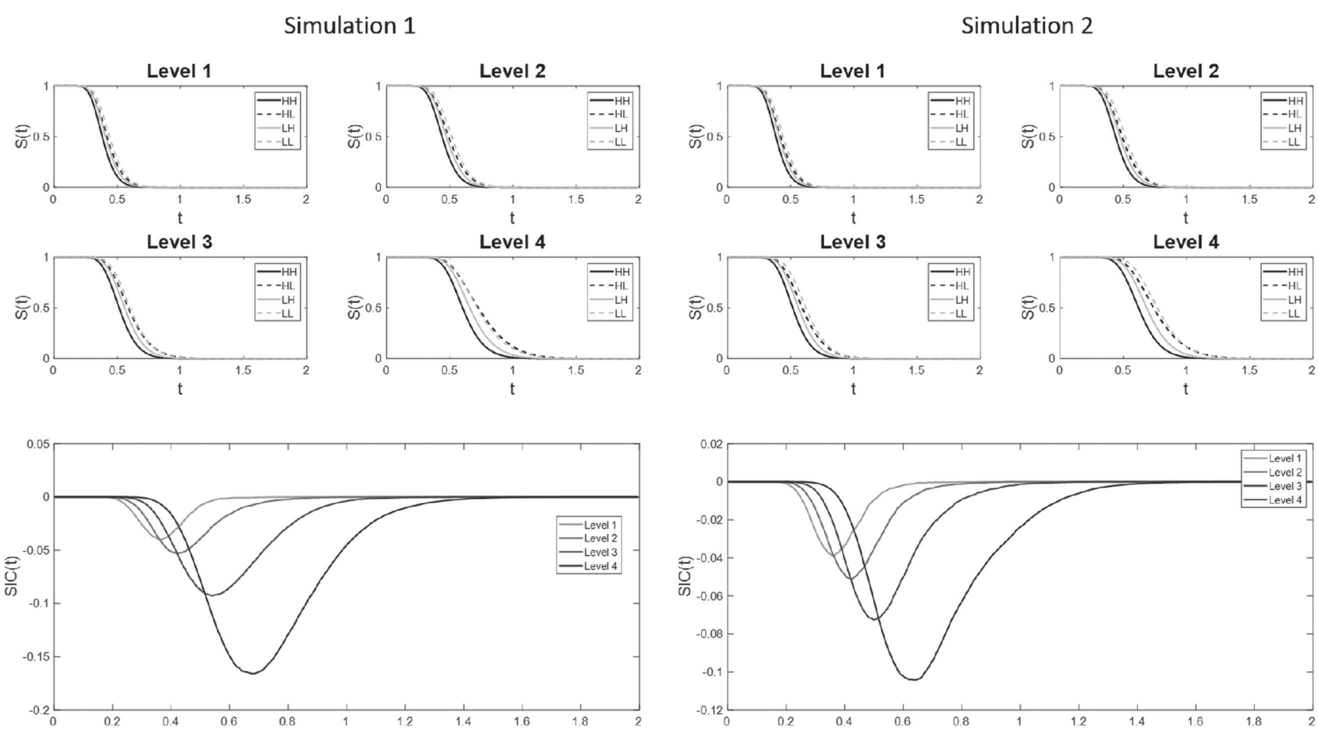
Note. RT = response time.

and one in which increasing salience increases both accumulators but the difference between the accumulators is constrained to be higher in the correct pair than the incorrect pair.

However, for the Poisson accumulator model, we additionally showed through formal proofs that the stochastic dominance relationship fails if the sum of the correct and error accumulators is constrained to be equal across high and low salience levels. Because the sum is constrained, increasing the rate of the correct accumulators in the correct channel necessarily decreases the rate of the error accumulators in the incorrect channel. This results in faster correct RTs for higher salience signals as expected but also faster error RTs for higher salience signals contrary to the usual stochastic dominance assumption. This by itself offers a strong qualitative test between two otherwise very similar race models.

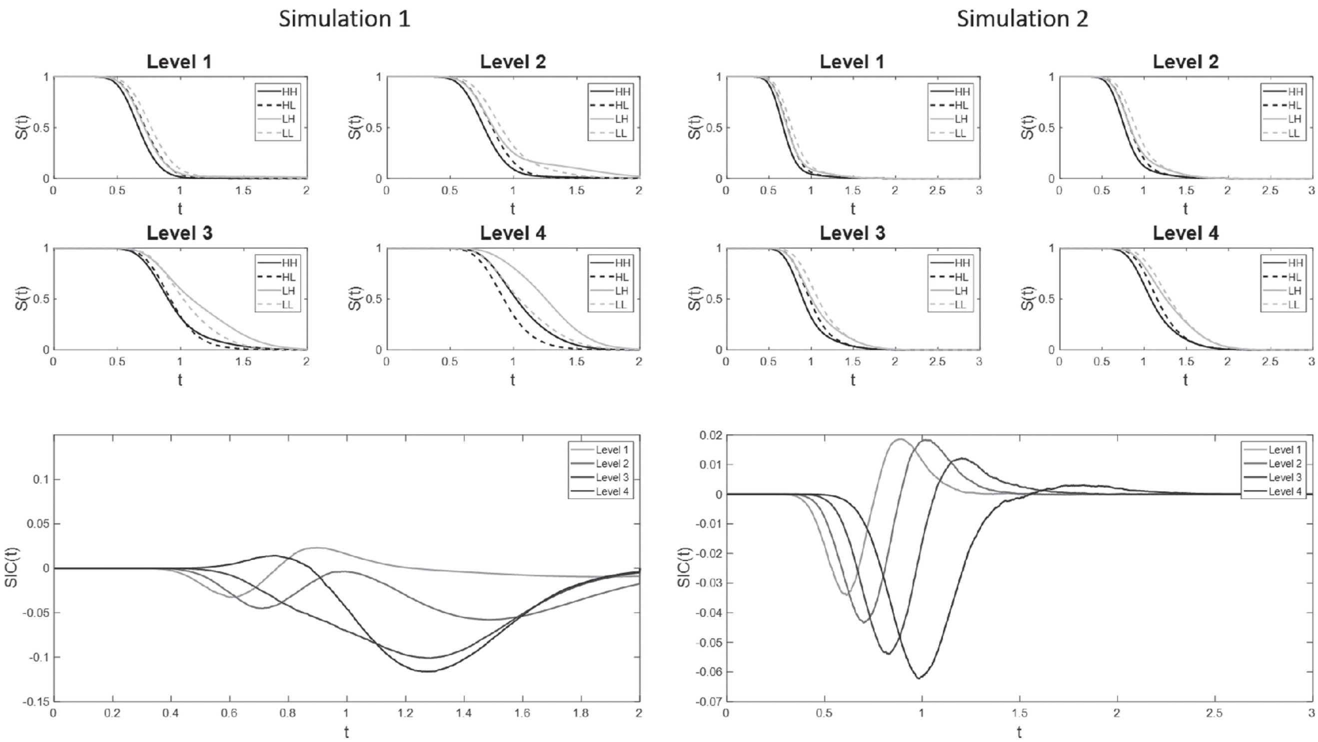
Using simulations we demonstrated that the shapes of the SICs hold when stochastic dominance is preserved but can vary when that assumption is violated. In practice, one should be able to confirm the stochastic ordering of the RTs, using visual inspection coupled with

Figure 12
Parallel Exhaustive OR Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

Figure 13
Serial Exhaustive OR Task Simulation Results



Note. The upper panels show the survivor functions conditioned on correct responding for the HH, HL, LH, and LL stimuli across different levels of parameter settings. The bottom panel shows the simulated SICs conditioned on accurate responding across each level of parameter settings. SIC = survivor interaction contrast.

statistical tests (Houpt et al., 2014), prior to examining the conditional SICs.

Robustness of SIC Predictions Under Errors and Other Approaches to Dealing With Error RTs

One important implication of the present results is that many previously published results, which typically do not have perfect accuracy, are upheld by the present analysis. These results have typically been accompanied by different arguments supporting the SIC analysis of RTs from items with imperfect accuracy. For example, Fific, Nosofsky, et al. (2008) provided simulations for exhaustive SICs demonstrating that "the serial and parallel architectures yield the expected MIC and SIC signatures even when error rates are very high (e.g., .30 for the LL stimulus)." (p. 360). In order to model the error RTs, Fific et al. introduced parametric instantiations of mental architecture models using random walks to account for error RTs distribution patterns from serial, parallel, and coactive

Table 9
Serial Exhaustive Mean RTs and Error Rates in an OR Task

	Level 1		Level 2		Level 3		Level 4	
	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc	Mean RT	Acc
				Simulation 1				
HH	678	1.00	790	1.00	945	.98	1,054	.75
HL	732	1.00	845	1.00	938	.92	943	.56
LH	745	1.00	943	1.00	1,138	.94	1,259	.58
LL	790	1.00	945	.98	1,055	.75	1,084	.26
				Simulation 2				
HH	692	1.00	802	1.00	945	.98	1,126	.93
HL	742	1.00	862	.99	1,013	.96	1,187	.86
LH	755	1.00	890	.99	1,071	.97	1,305	.87
LL	802	1.00	945	.98	1,126	.93	1,343	.75

Note. RT = response time.

models (see also Blunden et al., 2015, 2020; Cheng et al., 2018; Little et al., 2011, 2013; Moneer et al., 2016). Other theoretical treatments also support the robustness of canonical SIC predictions under low to moderate error rates (Townsend & Wenger, 2004). Our present work extends these results and places the argument for the incorporation of conditional RTs on much firmer theoretical ground.

Gondan (2019) recently investigated the incorporation of error RTs in the SIC using a different approach based on the use of censored survivor functions (i.e., Kaplan–Meier survival estimators). In estimating a censored survivor function, the assumption is that an error RT would have eventually been correct if given more time to respond. Censored survivor estimators treat errors as missing data in a principled way and recover a more accurate estimate of the survival function. Gondan also highlights that a general proof of the conditionally correct SIC (for the OR case) is not possible because of the possible violation of stochastic dominance. Here we refine this observation noting that the limitation is one of identifying the state of the model when the final response can arise from a mixture of states.

In the present work, we assume that errors are only caused by the process being stopped by the completion of an incorrect accumulator. We term this type of error an "accumulator failure". Bushmakin et al. (2017) recently argued that other types of processing errors were also possible. For instance, participants may simply fail to pay attention to the instructions or ignore (deliberately or not) the specified termination rule and self-terminate where exhaustive processing is required for accurate processing. The authors termed these types of errors "rule-breaking errors." A more general treatment of these types of errors may be necessary in order to determine whether the conditional SIC is useful in the case of rule-breaking. For instance, it may be possible to determine from the error SICs that rather than using a correct parallel exhaustive process, the rule-breaking error is better represented as a Serial Self-Termination. We leave this as a target for future work.

An inherent challenge in understanding and modeling error responses is the lack of one-to-one mapping between stimuli and responses in any underlying experimental design. That is, correct and error responses can arise in a variety of ways, and it is rarely evident what state led to a particular response. The SFT framework was developed in tandem with a theoretically driven methodology, the double factorial paradigm, which we have outlined at the beginning of this article and illustrated (partially) in Figures 1 and 2. Consider again the canonical example illustrated in these figures: Two signal positions could be occupied by two target signals at the same time, or one signal on the right, or one on the left, or no signals at all (and in addition, each presented signal could be of high or low salience, but that is tangential to the current discussion). In both the AND and OR tasks, multiple stimulus conditions are mapped onto the same response key. In the OR case, a correct "yes" response could reflect correct identification of both targets but could also be the consequence of correctly detecting only the right target while missing the left target, or vice versa. In fact, it could also be the outcome of numerous other errors, such as falsely reaching a "yes" decision in either channel when in fact no target was present. This issue means that there could be many ways for correct and incorrect responses to ensue, rendering analysis of correct and incorrect RTs, and the corresponding SIC difficult as our analyses show above.

To better understand and model error responses within the SFT framework, and perceptual decision-making in general, we have developed a novel task, the modified-AND-identification task, termed ID task for brevity (Howard et al., 2021). The new task is constructed in such a way that each possible stimulus type is mapped to a unique response option, and moreover a unique response key is associated with each of the assumed processing channels. For example, in a Canonical 2 (signal location: left, right) × 2 (signal present: yes, no) design, one response key corresponds to the channel assumed to process the left target, another to the right channel. Identifying two target signals at the same time requires to respond to both of them at the same time, by simultaneously pressing both keys within a window of 150 ms. And, finally, if no signal is presented, then an alternative third key should be pressed. The task provides improved traction on multi-accumulator models with behavioral data. Future work could adapt the task to deal with factorial manipulations of target salience, thereby providing the missing information about the response state for correct and error responses assisting with the estimation of the conditional SIC functions. In addition, the present investigation offers several important insights: (a) It shows what is possible or not from the point of view of accessing canonical SIC predictions conditional on accuracy or error. This pathway permits strong inferences as to architecture and decisional stopping rule when exhaustive processing takes place. (b) It investigated the two most popular and theoretically treated model classes, the DDM and the Poisson race model. While two of the race model instantiations follow the same dominance precepts as the diffusion model, the most popular makes dramatically different predictions as concerns the dominance relationships in errored data. (c) Finally, our investigation presents and discusses various ways that errors can occur. In particular, if errors of featural or dimensional identification take place, we considered and simulated ways in which model performance via SIC functions can go away from the canonical predictions.

References

Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of response times: Identification of cognitive mechanisms through mathematical modeling. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The oxford handbook of computational and mathematical psychology* (pp. 63–98). Oxford.

Algom, D., & Fitousi, D. (2016). Half a century of research on garner interference and the separability-integrality distinction. *Psychological Bulletin*, *142*(12), 1352–1383. https://doi.org/10.1037/bul0000072

Altieri, N., Fific, M., Little, D. R., & Yang, C.-T. (2017). A tutorial introduction and historical background to systems factorial technology. In D. R. Little, N. Aliteri, M. Fific, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms*. Elsevier.

Ashby, F. G., Boynton, G., & Lee, W. W. (1994). Categorization response times with multidimensional stimuli. *Perception and Psychophysics*, *55*(1), 11–27.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception and Psychophysics*, *6*, 321–326.

Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice 1. *British Journal of Mathematical and Statistical Psychology*, *18*, 207–225.

Blaha, L. M., & Houpt, J. W. (2015). An extension of workload capacity space for systems with more than two channels. *Journal of Mathematical Psychology*, *66*, 1–5.

Blunden, A. G., Hammond, D., Howe, P. D. L., & Little, D. R. (2021). Characterizing the time course of decision-making in change detection. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000306

Blunden, A. G., Howe, P. D. L., & Little, D. R. (2020). Evidence that within-dimension features are generally processed coactively. *Attention, Perception, and Psychophysics*, *82*(1), 193–227.

Blunden, A. G., Wang, T., Griffiths, D., & Little, D. (2015). Logical-rules and the classification of integral dimensions: Individual differences in the processing of arbitrary dimensions. *Frontiers in Psychology*, *5*, Article 310. https://doi.org/10.3389/fpsyg.2014.01531

Blurton, S. P., Kyllingsbæk, S., Nielsen, C. S., & Bundesen, C. (2020). A poisson random walk model of response times. *Psychological Review*, *127*, 362–411.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178.

Burns, D. M. (2016). Garner interference is not solely driven by stimulus uncertainty. *Psychonomic Bulletin and Review*, *23*(6), 1846–1853.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Bushmakin, M. A., Eidels, A., & Heathcote, A. (2017). Breaking the rules in perceptual information integration. *Cognitive Psychology*, *95*, 1–16.

Cheng, X. J., McCarthy, C. J., Wang, T. S., Palmeri, T. J., & Little, D. R. (2018). Composite faces are not (necessarily) processed coactively: A test using systems factorial technology and logical-rule models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *44*(6), 833–862.

Cox, G. E., & Criss, A. H. (2017). Parallel interactive retrieval of item and associative information from event memory. *Cognitive Psychology*, *97*, 31–61.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology*, *47*, 304–322.

Donders, F. C. (1868). On the speed of mental processes. *Acta Psychologica, Attention and Performance*, *II*, 20–54.

Donkin, C., Little, D. R., & Houpt, J. W. (2014). Assessing the speed–accuracy trade-off effect on the capacity of information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1183–1202.

Fific, M., Little, D. R., & Nosofsky, R. (2010). Logical-rule models of classification response times: A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review*, *117*, 309–348.

Fific, M., Nosofsky, R. M., & Townsend, J. (2008). Information-processing architectures in multidimensional classification: A validation of test of the systems factorial technology. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 356–375.

Fific, M., Townsend, J., & Eidels, A. (2008). Studying visual search using systems factorial methodology with target-distractor similarity as the factor. *Attention, Perception, and Psychophysics*, *70*, 583–603.

Fific, M., & Townsend, J. T. (2010). Information-processing alternatives to holistic perception: Identifying the mechanisms of secondary-level holism within a categorization paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), Article 1290.

Forstmann, B. U., Wagenmakers, E.-J. (Eds.). (2015). *An introduction to model-based cognitive neuroscience*. Springer.

Garner, W. R. (1974). *The processing of information and structure*. Psychology Press.

Garrett, P. M., Howard, Z., Houpt, J. W., Landy, D., & Eidels, A. (2019). Comparative estimation systems perform under severely limited workload capacity. *Journal of Mathematical Psychology*, *92*(3), Article 102255.

Gondan, M. (2019). Incorrect responses in the response time interaction contrast. *Journal of Mathematical Psychology*, *92*(1–2), Article 102249.

Greeno, J. G., & Steiner, T. E. (1964). Markovian processes with identifiable states: General considerations and application to all-or-none learning. *Psychometrika*, *29*, 309–333.

Griffiths, D. W., Blunden, A. G., & Little, D. R. (2017). Logical-rule based models of categorization: Using systems factorial technology to understand feature and dimensional processing. In D. R. Little, N. Aliteri, M. Fific, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms* (pp. 245–269). Academic Press.

Harding, B., Goulet, M.-A., Jolin, S., Tremblay, C., Villeneuve, S.-P., & Durand, G. (2016). Systems factorial technology explained to humans. *Tutorials in Quantitative Methods for Psychology*, *12*(1), 39–56.

Heathcote, A., Wagenmakers, E.-J., & Brown, S. D. (2014). The falsifiability of actual decision-making models. *Psychological Review*, *121*, 676–678.

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*, 307–330.

Houpt, J. W., Eidels, A., & Little, D. R. (2019). Developments in systems factorial technology: Theory and applications. *Journal of Mathematical Psychology*, *92*, 1–3.

Houpt, J. W., & Little, D. R. (2017). Statistical analysis of the resilience function. *Behavior Research Methods*, *49*, 1261–1277.

Howard, Z. L., Belevski, B., Eidels, A., & Dennis, S. (2020). What do cows drink? A systems factorial technology account of processing architecture in memory intersection problems. *Cognition*, *202*, Article 104294.

Howard, Z. L., Garrett, P., Little, D. R., Townsend, J. T., & Eidels, A. (2021). A show about nothing: No-signal processes in systems factorial technology. *Psychological Review*, *128*(1), 187–201.

Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, *121*, 1–32.

Khodadadi, A., & Townsend, J. T. (2015). On mimicry among sequential sampling models. *Journal of Mathematical Psychology*, *68*, 37–48.

LaBerge, D. (1994). Quantitative models of attention and response processes in shape identification tasks. *Journal of Mathematical Psychology*, *38*, 198–243.

Laming, D. R. J. (1968). *Information theory of choice reaction times*. Wiley.

Link, S. W. (1992). *The wave theory of difference and similarity*. Psychology Press.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Pscyhometrika*, *40*, 77–105.

Little, D. R., Altieri, N., Fific, M., & Yang, C.-T. (2017). *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms*. Academic Press.

Little, D. R., Eidels, A., Fific, M., & Wang, T. (2015). Understanding the influence of distractors on workload capacity. *Journal of Mathematical Psychology*, *68*, 25–36.

Little, D. R., Eidels, A., Fific, M., & Wang, T. (2018). How do information processing systems deal with conflicting information? Differential predictions for serial, parallel and coactive processing models. *Computational Brain and Behavior*, *1*, 1–21.

Little, D. R., Eidels, A., Houpt, J. W., & Yang, C. T. (2017). Set size slope still does not distinguish parallel from serial search. *Behavioral and Brain Sciences*, *40*, 32–33.

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

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

Little, D. R., Wang, T., & Nosofsky, R. M. (2016). Sequence-sensitive exemplar and decision-bound accounts of speeded-classification performance in a modified garner-tasks paradigm. *Cognitive Psychology*, *89*, 1–38.

Logan, G. D. (1996). The code theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103*, Article 603.

Logan, G. D. (2002). Parallel and serial processing. *Stevens' handbook of experimental psychology: Methodology in experimental psychology* (pp. 271–300). Wiley.

Lowe, K. A., Reppert, T. R., & Schall, J. D. (2019). Selective influence and sequential operations: A research strategy for visual search. *Visual Cognition*, *27*, 387–415.

Luce, R. D. (1963). *Detection and recognition* [Newspaper article]. Wiley.

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

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

Moneer, S., Wang, T., & Little, D. R. (2016). The processing architectures of whole-object features: A logical rules approach. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1443–1465.

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review*, *116*(3), 499–518.

Nosofsky, R. M., & Palmeri, T. (1997a). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception and Psychophysics*, *59*, 1027–1048.

Nosofsky, R. M., & Palmeri, T. (1997b). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Pike, A. R. (1966). Stochastic models of choice behavior: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, *19*, 15–32.

Pike, R. (1973). Response latency models for signal detection. *Psychological Review*, *80*(1), 53–68.

Raab, D. (1962). Statistical facilitation of simple reaction time. *Transaction of the New York Academy of Science*, *43*, 574–590.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *5*, 347–356.

Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 127–140.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260–281.

Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, *106*, 261–300.

Roberts, F. S. (1979). *Measurement theory with applications to decision making utility and the social sciences*. Addison-Wesley.

Schweickert, R. (1978). A critical path generalization of the additive factor methods analysis of a Stroop task. *Journal of Mathematical Psychology*, *18*, 105–139.

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

Shang, L., Little, D. R., Webb, M. E., Eidels, A., & Yang, C.-T. (2021). The workload capacity of semantic search in convergent thinking. *Journal of Experimental Psychology: General*, *150*(11), 2230–2245. https://doi.org/10.1037/xge0001045

Shepard, R. N., & Chang, J.-J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, *65*(1), 94–102.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42.

Smith, P. L. (2010). From poisson shot noise to the integrated Ornstein–Uhlenbeck process: Neurally principled mode of information accumulation in decision-making and response time. *Journal of Mathematical Psychology*, *54*(2), 266–283.

Smith, P. L., Ratcliff, R., & McKoon, G. (2014). The diffusion model is not a deterministic growth model: Comment on Jones and Dzhafarov (2014). *Psychological Review*, *121*, 679–688.

Smith, P. L., & Van Zandt, T. (2000). Time-dependent poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, *53*, 293–315.

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

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

Tillman, G., Van Zandt, T., & Logan, G. D. (2020). Sequential sampling models without random between-trial variability: The racing diffusion model of speeded decision making. *Psychonomic Bulletin and Review*, *27*, 911–936.

Townsend, J. T. (1971a). A note on the indentifiability of parallel and serial processes. *Perception and Psychophysics*, *10*, 161–163.

Townsend, J. T. (1971b). Theoretical analysis of an alphabetic confusion matrix. *Perception and Psychophysics*, *9*, 40–50.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, *28*, 363–400.

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). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, *108*, 551–567.

Townsend, J. T. (1992). On the proper scales for reaction time. In H. Geissler, S. Link, & J. T. Townsend (Eds.), *Cognition, information processing, and psychophysics: Basic issues* (pp. 105–120). Lawrence Erlbaum.

Townsend, J. T., & Altieri, N. (2012). An accuracy-response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review*, *119*(3), 500–516.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In M. J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Lawrence Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge University Press.

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

Townsend, J. T., & Landon, D. E. (1983). Mathematical models of recognition and confusion in psychology. *Mathematical Social Sciences*, *4*, 25–71.

Townsend, J. T., & Liu, Y. (2020). Can the wrong horse win: The ability of race models to predict fast or slow errors. *Journal of Mathematical Psychology*, *97*, Article 102360.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*, *39*, 321–340.

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*, 1003–1035.

Townsend, J. T., Wenger, M. J., & Houpt, J. W. (2018). Uncovering mental architecture and related mechanisms in elementary human perception, cognition, and action. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*, *5*, 1–30.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin and Review*, *7*, 424–465.
Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin and Review*, *7*, 208–256.
Van Zandt, T., & Ratcliff, R. (1995). Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomics Bulletin and Review*, *2*, 20–54.
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 and Performance*, *39*(5), 1365–1385.
Yang, C.-T., Hsu, H.-Y., & Yeh, Y.-Y. (2011). Relative salience affects the process of detecting changes in orientation and luminance. *Acta Psychologica*, *138*, 377–389.
Yang, C.-T., Little, D. R., & Hsu, C.-C. (2014). The influence of cuing on attentional focus in perceptual decision making. *Attention, Perception and Psychophysics*, *76*(8), 2256–2275.
Yang, H., Fific, M., & Townsend, J. T. (2014). Survivor interaction contrast wiggle predictions of parallel and serial models for an arbitrary number of processes. *Journal of Mathematical Psychology*, *59*, 82–94.
Yang, H., Little, D. R., Eidels, A., & Townsend, J. T. (2020, May). *Survivor interaction contrasts for errored response times: Non-parametric contrasts for serial and parallel systems*. PsyArXiv. https://doi.org/10.31234/osf.io/jtk6q

Appendix A

Diffusion Model Proofs

Proposition 9. Given two Wiener diffusion models with different drift rates, v_L and v_H, such that $0 \leq v_L \leq v_H$, then the likelihood ratio function, $L(t) = f_H(t|R_a)/f_L(t|R_a)$, is monotonically decreasing with respect to the value of t. Consequently, the CDF for H stochastically dominates the CDF for L; that is, $F_H(t|R_a) > F_L(t|R_a)$ for all t.

Proof. The likelihood ratio function of two Wiener diffusion models can be written as:

$$L(t) = \frac{P_L(R_a)}{P_H(R_a)} \times \exp\left[\left(\frac{\nu_H(a-z)}{s^2} - \frac{\nu_H^2 t}{2s^2}\right) - \left(\frac{\nu_L(a-z)}{s^2} - \frac{\nu_L^2 t}{2s^2}\right)\right]$$
$$= C \times \exp\left[\left(\frac{\nu_L^2}{2s^2} - \frac{\nu_H^2}{2s^2}\right)t\right].$$

Where $C = P_L(R_a)/P_H(R_a) \times \exp[\frac{(\nu_H-\nu_L)(a-z)}{s^2}]$ is a positive constant.

Since $v_L \leq v_H$, $L'(t) < 0$, the exponential term will decrease with t (i.e., be increasingly negative), which implies that $L(t)$ is a decreasing function t. □

Corollary 1. *The likelihood ratio function of two Wiener diffusion models with rates* $v_L \leq v_H$, $L(t) = f_H(t|R_b)/f_L(t|R_b)$, *will be monotonically decreasing in t. This implies that* $F_H(t|R_b) > f_L(t|R_b)$ for all t.

Proof. The proof is the same as above replacing $P(R_a)$ with $P(R_b)$ and $a - z$ with $b - z$. □

Proposition 10. Assume that we have two distributions, one for H and one for L, with densities that possess a parameter v which may be interpreted as a random variable (V), plus other parameters that are fixed and the same. Although the next assumption is easily dropped, we can think of these distributions involving v as also containing the variable t which represents processing time as in $f(t; v)$. Next, suppose that we introduce a third density on that (random) parameter V; call it $q(v; u)\, dv$, where u is a parameter that orders the cumulative distribution such that $Q(v; u_L) = P(V \leq v; u_L) < Q(v; u_H) = P(V \leq v; u_H)$ if and only if $u_L < u_H$. For simplicity, we also assume the distribution ordering comes about through a single-point crossover of the densities $q(v; u_H)$ and $q(v; u_L)$. In the context of SFT, we can think of u as representing an entity that can be influenced by an experimental factor.

Then, the resultant distributions on t, deriving from the respective probability mixtures, are ordered as $\int_{-\infty}^{\infty} F(t;\nu)q(\nu;u(H)d\nu) > \int_{-\infty}^{\infty} F(t;\nu)q(\nu;u(L)d\nu)$.

Proof. Without loss of generality, we presume that the support of V is the real line. Because F is always growing in v for all values of t, and the positive area associated with $q(v; u_H) - q(v; u_L)$ is equal to the entire negative area we see that the above inequality has to hold. □

Our original article completing the distributional dominance hierarchy indicated that actual data indicated that single-point crossovers of densities might not be an anomaly and considerably stronger dominance, such as through an order on hazard functions or a monotonic likelihood function, were likewise found in some standard short-term memory search data (Townsend, 1990b).

Corollary 2. *Set* $F(t; v)$, *plus the remaining constant parameters, to be the Wiener diffusion model with drift v and consider any probability mixture obeying the above stricture; the ordering follows. A special case of interest is found when q is Gaussian with mean u.*

Proof. When the variance is the same in the two distributions, H and L, the proper ordering results due to the fact that $u_L < u_H$. □

Proposition 11. Suppose two cumulative distributions functions are ordered in some variable, such as time t, to give $F(t; v, u_H) > F(t; v, u_L)$. Also, suppose that they depend on a second random variable, for instance, Z. Then, the probability mixture of the two distribution over $Z = z$ will be ordered in the same direction.

(*Appendices continue*)

Proof. Assume that Z is supported over an interval $[a, b]$. Integrating $\int_a^b [F(t;\nu,u_H) - F(t;\nu,u_L)] \cdot p(z)dz$ demonstrates that the average difference is positive. ◻

Corollary 3. *The probability mixture over the drift rate yields, from Proposition 10 an ordering in u. Hence, an ensuing mixture of these two distributions over a starting point, now called z, is also ordered.*

Corollary 4. *Given two Wiener diffusion models with different drift rates, v_L and v_H, such that $0 \le v_L \le v_H$, and a starting point, z, distributed according to any density, p(z), on the domain of z the compound distribution for H stochastically dominates the compound distribution for L, that is, $\bar{F}_H(t|R_b) > \bar{F}_L(t|R_b)$ for all t.*

Proof. The proof follows from Theorem 3 where the density over v is a point density with 0 variance. ◻

Appendix B

Poisson Counter Model Proofs

Proposition 12. Consider two Poisson counter models with different rates, one of high salience, v_H^C, v_I^H, and one of low salience, v_C^L, v_I^L. Assume $v_C^L < v_C^H$ and $v_C^H + v_I^H = v_C^L + v_I^L$, then $L(t) = g_H(t|R_a)/g_H(t|R_a)$ is monotonically decreasing with respect to the value of t. Thus, $G_H(t|R_a) > G_L(t|R_a)$ for all t.

To prove Proposition 12, we require the following lemma:

Lemma 1. For any positive integer K, $L_K(t) = [\sum_{j=0}^K (ct)^j/j!]/[\sum_{j=0}^K t^j/j!]$ is decreasing with t when $c < 1$, and increasing with t when $c > 1$.

Proof. The Lemma is proved by induction on K, first for $c < 1$. When $K = 1$,

$$L_1(t) = \frac{1 + ct}{1 + t},$$

and the derivative of $L(t)$ is:

$$L1(t) = \frac{c - 1}{(1 + t)^2},$$

which is negative since $c < 1$. Thus, $L(t)$ is decreasing when $K = 1$. Now, given the induction hypothesis that $L(t)$ is decreasing when $K = k$, we prove that this is true for $K = k + 1$. To begin, we write $L_K(t)$ as:

$$L_K(t) = \frac{P_K(ct)}{P_K(t)},$$

where:

$$P_K(t) = \sum_{j=0}^{K} [(t)^j/j!].$$

The derivative of $L_K(t)$ is:

$$LK(t) = \frac{[P_K(ct)]' P_K(t) - P_K(ct)PK(t)}{[P_K(t)]^2}.$$

Note that

$$d[P_K(t)]/dt = P_K(t) - \frac{t^K}{K!} = P_{K-1}(t),$$

and

$$d[P_K(ct)]/dt = cPK(ct) = c\left[P_K(ct) - \frac{(ct)^K}{K!}\right] = cP_{K-1}(ct).$$

Thus, the derivative can be written as:

$$LK(t) = \frac{cP_{K-1}(ct)P_K(t) - P_K(ct)P_{K-1}(t)}{[P_K(t)]^2} < 0.$$

Therefore, we have by the induction hypothesis that the numerator of the proceeding is negative $K = k$, that is,

$$N_k(t) \triangleq cP_{K-1}(ct)P_K(t) - P_K(ct)P_{K-1}(t) < 0.$$

For $K = k + 1$, we have:

$$\begin{aligned} N_{k+1}(t) &= c \times \left[P_{k-1}(ct) + \frac{c^k t^k}{k!}\right]\left[P_k(t) + \frac{t^{k+1}}{(k+1)!}\right] \\ &\quad - \left[P_k(ct) + \frac{c^{k+1}t^{k+1}}{(k+1)!}\right]\left[P_{k-1}(t) + \frac{t^k}{k!}\right] \\ &= N_k(t) + \frac{t^k}{k!}\left[c^{k+1}P_k(t) + \frac{ct}{(k+1)}P_{k-1}(ct) \right. \\ &\quad \left. - \frac{c^{k+1}t}{k+1}P_{k-1}(t) - P_k(ct)\right]. \end{aligned}$$

Rearranging, we can write:

$$\begin{aligned} N_{k+1}(t) &= N_k(t) + \frac{t^k}{k!}\left[c^{k+1} - 1 + [c^{k+1} - c]\left[1 - \frac{1}{k+1}\right]\right. \\ &\quad \left. + \sum_{i=2}^{k}[c^{k+1} - c^i]\frac{1}{(i-1)!}\left[\frac{1}{i} - \frac{1}{k+1}\right]\right]. \end{aligned} \tag{B1}$$

Since $c < 1$ and $N_k(t) < 0$, it is easy to see that $N_{k+1}(t) < 0$, therefore $L_{k+1}(t)$ is decreasing with t. By induction, the proposition is proved for $c < 1$.

The proof for $c > 1$ proceeds in an analogous fashion, so we just point out the difference here: When $k = 1$, $L'(t)$ is positive which implies that $L(t)$ is increasing with t. Based on Equation 7 and the fact that $c > 1$, when $L_k(t)$ is increasing, we can prove that $L_{k+1}(t)$ is also increasing. Therefore, we have proved that $c > 1$. ◻

Now we are ready for the proof of Proposition 12.

(*Appendices continue*)

Proof. The likelihood ratio of the low to high salience Poisson race model is

$$L(t) = \frac{P_L(a)}{P_H(a)} \frac{(v_C^H t)^{K_a - 1} v_C^H \exp(-v_C^H t)}{(v_C^L t)^{K_a - 1} v_C^L \exp(-v_C^L t)} \frac{\exp(-v_I^H t)}{\exp(-v_I^L t)} \frac{\sum_{j=0}^{K_b - 1} (v_I^H t)^j / j!}{\sum_{j=0}^{K_b - 1} (v_I^L t)^j / j!}.$$

Since $v_C^H + v_I^H = v_C^L + v_I^L$, we can simplify the equation as follows:

$$L(t) = \frac{P_L(a)}{P_H(a)} \left(\frac{v_C^H}{v_C^L}\right)^{K_a} \times \frac{\sum_{j=0}^{K_b - 1} (v_I^H t)^j / j!}{\sum_{j=0}^{K_b - 1} (v_I^L t)^j / j!}.$$

Let $v_I^L t = \tau$, then

$$L(\tau) = C \times \frac{\sum_{j=0}^{K_b - 1} (c\tau)^j / j!}{\sum_{j=0}^{K_b - 1} (\tau)^j / j!},$$

in which

$$C = \frac{P_L(a)}{P_H(a)} \left(\frac{v_C^H}{v_C^L}\right)^{K_a},$$

and $c = v_I^H / v_I^L$ is a constant smaller than 1. Therefore, it follows from Lemma 1 that $L(\tau)$ is a decreasing function, which finishes the proof. □

Proposition 13. Consider two Poisson counter models with different rates, one of high salience, v_C^H, v_I^H, and one of low salience, v_C^L, v_I^L. Assume $v_C^L < v_C^H$ and $v_C^H + v_I^H = v_C^L + v_I^L$, then $L(t) = g_H(t|R_b)/g_L(t|R_b)$ is monotonically increasing with respect to the value of t. Thus, $G_H(t|R_b) > G_L(t|R_b)$ for all t.

Proof. The likelihood ratio of the low to high salience Poisson race model is:

$$L(t) = \frac{P_L(b)}{P_H(b)} \frac{(v_I^H t)^{K_b - 1} v_I^H \exp(-v_I^H t)}{(v_I^L t)^{K_b - 1} v_I^L \exp(-v_I^L t)} \times \frac{\exp - v_C^H(t)}{\exp - v_C^L(t)} \frac{\sum_{j=0}^{K_a - 1} (v_C^H t)^j / j!}{\sum_{j=0}^{K_a - 1} (v_C^L t)^j / j!}.$$

As in the previous proof, we can simplify $L(t)$ as follows:

$$L(\tau) = C \times \frac{\sum_{j=0}^{K_a - 1} (c\tau)^j / j!}{\sum_{j=0}^{K_a - 1} (\tau)^j / j!},$$

in which

$$C = \frac{P_L(b)}{P_H(b)} \left(\frac{v_I^H}{v_I^L}\right)^{K_b},$$

where $v_C^L t = \tau$ and $c = v_C^H / v_C^L$ is a constant greater than 1. Following Lemma 1, $L(\tau)$ is an increasing function, which finishes the proof. □

Proposition 14. Consider two Poisson counter models with different rates, v_C^H, v_I^H and v_C^L, v_I^L. Assume $v_C^L < v_C^H$ and $v_I^H = v_I^L$, then $L(t) = g_H(t|R_a)/g_H(t|R_a)$ is monotonically decreasing with respect to the value of t. Thus, $G_H(t|R_a) > G_L(t|R_a)$ for all t.

Proof. We have be Equation 19 and the assumption that $v_I^H = v_I^L$ that

$$L(t) = \frac{P_L(a)}{P_H(a)} \frac{(v_C^H)^{K_a}}{(v_C^L)^{K_a}} \exp\left[(v_C^L - v_C^H)t\right].$$

The first two terms on the right are positive; since $v_C^L < v_C^H$, $L'(t)$ is negative, and $L(t)$ is a decreasing function of t. □

Proposition 15. Consider two Poisson counter models with different rates, v_C^H, v_I^H and v_C^L, v_I^L. Assume $v_C^L < v_C^H$ and $v_I^H = v_I^L$, then $L(t) = g_H(t|R_b)/g_H(t|R_b)$ is monotonically decreasing with respect to the value of t. Thus, $G_H(t|R_b) > G_L(t|R_b)$ for all t.

Proof. By Equation 20 and the assumption that $v_I^H = v_I^L$, we have

$$L(t) = \frac{P_L(b)}{P_H(b)} \frac{\exp(v_C^L t)}{\exp(v_C^H t)} \frac{\sum_{j=0}^{K_a - 1} (v_C^H t)^j / j!}{\sum_{j=0}^{K_a - 1} (v_C^L t)^j / j!}.$$

Let $C = P_L(b)/P_H(b)$, $n_L(t) = \exp(v_C^L t)$, $n_H(t) = \exp(v_C^H t)$, $m_H(t) = \sum_{j=0}^{K_a - 1} (v_C^H t)^j / j!$, and $m_L(t) = \sum_{j=0}^{K_a - 1} (v_C^L t)^j / j!$. We can rewrite $L(t)$ as:

$$L(t) = C \times \frac{n_L(t)}{n_H(t)} \frac{m_H(t)}{m_L(t)}.$$

Hence, the statement that $L(t)$ is monotonically decreasing is equivalent to:

$$\left[\frac{n_L(t)}{n_H(t)} \frac{m_H(t)}{m_L(t)}\right]' = \left[\frac{n_L(t)}{n_H(t)}\right]' \frac{m_H(t)}{m_L(t)} + \frac{n_L(t)}{n_H(t)} \left[\frac{m_H(t)}{m_L(t)}\right]' < 0.$$

Since

$$\left[\frac{n_L(t)}{n_H(t)}\right]' = \left[\exp\left[(v_C^L - v_C^H)t\right]\right]' = (v_C^L - v_C^H) \frac{n_L(t)}{n_H(t)},$$

the preceding inequality is equivalent to:

$$v_C^L - v_C^H + \frac{m_H'(t)}{m_H(t)} - \frac{m_L'(t)}{m_L(t)} < 0.$$

We also note that:

$$m'(t) = v_C \sum_{j=0}^{K_a - 2} \frac{(v_C t)^j}{j!}.$$

Substituting, we have:

$$\frac{(v_C^L)^{K_a}}{\sum_{j=0}^{K_a - 1} (v_C^L t)^j / j!} - \frac{(v_C^H)^{K_a}}{\sum_{j=0}^{K_a - 1} (v_C^H t)^j / j!} < 0.$$

It is easy to prove that an equation of the form $\frac{(v)^k}{\sum_{j=0}^{k-1} (vt)^j / j!}$ is monotonically increasing with v when $v > 0$, thus the proceeding inequality holds. □

(*Appendices continue*)

Proposition 16. Consider two Poisson counter models with different rates, ν_C^H, ν_I^H and ν_C^L, ν_I^L. Assume $\nu_C^L < \nu_C^H$ and $\nu_I^L < \nu_I^H$, and $\nu_C^H - \nu_C^L > \nu_I^H - \nu_I^L$, then $L(t) = g_H(t|R_a)/g_H(t|R_a)$ is monotonically decreasing with respect to the value of t. Thus, $G_H(t|R_a) > G_L(t|R_a)$ for all t.

Proof. Given the Poisson density function, $L(t)$ can be expressed as

$$L(t) = \exp[(\nu_C^L + \nu_I^L - \nu_C^H - \nu_I^H)t] \times \frac{\sum_{j=0}^{K_b-1} (\nu_I^H t)^j / j!}{\sum_{j=0}^{K_b-1} (\nu_I^L t)^j / j!} \times C,$$

in which

$$C = \frac{P_L(a)}{P_H(a)} \left(\frac{\nu_C^H}{\nu_C^L}\right)^{K_a}.$$

By the assumption that $\nu_C^L + \nu_I^L - \nu_C^H - \nu_I^H < 0$, the first term in $L(t)$ is decreasing with t. Since $\nu_I^H > \nu_I^L$, we have by the proof of Proposition 12, that the second term is also decreasing. Combined with the fact that C is positive, we can conclude that $L(t)$ is monotonically decreasing with t. □

Proposition 17. Consider two Poisson counter models with different rates, ν_C^H, ν_I^H and ν_C^L, ν_I^L. Assume $\nu_C^L < \nu_C^H$ and $\nu_I^L < \nu_I^H$, and $\nu_C^H - \nu_C^L > \nu_I^H - \nu_I^L$, then $L(t) = g_H(t|R_b)/g_H(t|R_b)$ is monotonically decreasing with respect to the value of t. Thus, $G_H(t|R_b) > G_L(t|R_b)$ for all t.

Proof. Given Equation 20, we can write the likelihood ratio as

$$L(t) = \exp[(\nu_C^L - \nu_C^H)t] \times \frac{\sum_{j=0}^{K_a-1} (\nu_C^H t)^j / j!}{\sum_{j=0}^{K_a-1} (\nu_C^L t)^j / j!} \times C(t),$$

where

$$C(t) = \exp[(\nu_C^L - \nu_C^H)t] \frac{P_L(b)}{P_H(b)} \left(\frac{\nu_I^H}{\nu_I^L}\right)^{K_b}.$$

Since $\nu_I^L - \nu_I^H < 0$, $C(t)$ is decreasing with t. Following the proof of Proposition 15, we can see that the ratio

$$\frac{\exp(\nu_C^L t)}{\exp(\nu_C^H t)} \frac{\sum_{j=0}^{K_a-1} (\nu_C^H t)^j / j!}{\sum_{j=0}^{K_a-1} (\nu_C^L t)^j / j!},$$

is also decreasing with t. Therefore, $L(t)$ is a decreasing function of t. □

Received May 13, 2020
Revision received December 16, 2021
Accepted December 19, 2021 ■