Journal of Mathematical Psychology 56 (2012) 476–494

Contents lists available at SciVerse ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

General recognition theory extended to include response times: Predictions for a class of parallel systems

James T. Townsend [a], Joseph W. Houpt [b,*], Noah H. Silbert [c]

[a] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA

[b] Department of Psychology, Wright State University, Dayton, OH 45435, USA

[c] Center for Advanced Study of Language, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 21 February 2011
Received in revised form
28 August 2012
Available online 22 November 2012

Keywords:
General recognition theory
Perceptual separability
Perceptual independence
Response times
Linear stochastic systems
Parallel processing
Cognitive architecture

ABSTRACT

General Recognition Theory (GRT; Ashby & Townsend, 1986) is a multidimensional theory of classification. Originally developed to study various types of perceptual independence, it has also been widely employed in diverse cognitive venues, such as categorization. The initial theory and applications have been static, that is, lacking a time variable and focusing on patterns of responses, such as confusion matrices. Ashby proposed a parallel, dynamic stochastic version of GRT with application to perceptual independence based on discrete linear systems theory with imposed noise (Ashby, 1989). The current study again focuses on cognitive/perceptual independence within an identification classification paradigm. We extend stochastic GRT and its implicated methodology for cognitive/perceptual independence, to an entire class of parallel systems. This goal is met in a distribution-free manner and includes all linear and non-linear systems satisfying very general conditions. A number of theorems are proven concerning stochastic forms of independence. However, the theorems all assume the stochastic version of decisional separability. A vital task remains to investigate the consequences of failures of stochastic decisional separability.

1. Introduction

Essentially every sensory stimulus a person encounters on a day to day basis can be characterized, at least in part, in terms of values on multiple physical and psychological dimensions. Even the simplest visual input consists of particular values of hue, saturation, brightness, and shape, and even the simplest auditory input has a characteristic amplitude envelope and frequency. Analogous properties can be defined for input to other sensory systems. The mapping between any given physical dimension and a corresponding psychological scale is, of course, the core of psychophysics. But perception is not exhaustively described by cataloging and analyzing such one-to-one mappings. Multidimensional physical structure requires that we also consider the relationships between psychological dimensions and the potentially very complex mapping between multiple physical and multiple psychological dimensions.

General Recognition Theory (GRT) was developed, in large part, for the purpose of modeling the relationships between psychological dimensions. In its basic form, GRT is, like signal detection theory (SDT; Green & Swets, 1966; MacMillan & Creelman, 1991), a two-stage model of perception and response selection. In GRT, as in SDT, perception is assumed to be noisy and partially random while response selection is assumed to be deterministic. Repeated exposure to multidimensional stimuli produces multivariate perceptual distributions in a perceptual space partitioned by multidimensional response criteria. These perceptual and decisional processes enable the relationships between psychological dimensions to be characterized in terms of perceptual and decisional dependencies in GRT (e.g., Ashby & Townsend, 1986; Cornes, Donnelly, Godwin, & Wenger, 2011; Farris, Treat, & Viken, 2010; Kadlec & Townsend, 1992a; Maddox, 1992; Richler, Gauthier, Wenger, & Palmeri, 2008). Variants of the basic GRT structure—noisy perception and deterministic response selection in multiple dimensions—have also been used to represent and analyze perceptual similarity (e.g., Ashby & Perrin, 1988) as well as providing the foundation for the Decision Bound theory of categorization (e.g., Ashby & Gott, 1988; Ashby & Maddox, 1990, 1992, 1993; Maddox, 1995; Maddox & Bohil, 1998).

Work in the GRT framework that is focused on relationships between dimensions most often employs a closely aligned experimental paradigm known as “feature-complete, full-factorial identification”. The basic idea of this experimental paradigm is that stimuli are defined in terms of the exhaustive factorial combination of multidimensional attributes (for a set of attributes of interest). To make this more concrete, suppose, for example, that we are

* Corresponding author.
E-mail addresses: jtownsen@indiana.edu (J.T. Townsend), joseph.houpt@wright.edu (J.W. Houpt), nsilbert@umd.edu (N.H. Silbert).

0022-2496/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmp.2012.09.001

interested in studying visual dependences in the perception of hue and shape. We might define a set of stimuli in terms of a factorial combination of the hues red and purple and the shapes square and rectangle. A full-factorial combination of these features would give us four stimuli: red square, purple square, red rectangle, and purple rectangle. In an identification task, each of these stimuli has a unique response associated with it. Of course, other tasks can also use factorially defined stimuli, as well. For example, a same-different discrimination task could map these stimuli onto a smaller number of responses (e.g., Thomas, 1996), and a "concurrent ratings" task could require a larger number of responses, requiring multiple levels of confidence ratings on each dimension (e.g., for the presence of rectangles or purple hue; see, e.g., Ashby, 1988).

As with SDT, the original GRT model was static, providing an account of confusion count data in the full-factorial identification paradigm. Most ensuing theoretical efforts and empirical applications have also been static. Nonetheless, the static GRT model has a broad purview, encompassing non-parametric and distribution free tests of independence as well as more fine-grained models employing multivariate Gaussian perceptual distributions (e.g., Böckenholt, 1992; Ennis & Mullen, 1992; Kadlec & Townsend, 1992a; Marley, 1992; Silbert, 2012; Soete & Carroll, 1992; Thomas, 1999; Wickens & Olzak, 1992). We seek, in the present work, to extend "full-factorial" GRT by combining patterns of (in)accurate identifications and confusions with response times (RTs). A number of prominent parameterized models which include both RT and accuracy have been available for some time (e.g., Laming, 1968; Link & Heath, 1975; Ratcliff, 1978; Smith & Van Zandt, 2000; Smith & Vickers, 1988; Townsend & Ashby, 1983), though there has been little effort in applying these to the study of perceptual dependencies.

Discrete and continuous time linear systems have been used to take important steps in bringing stochasticity to GRT in the past (Ashby, 1989, 2000), though even here the focus has been on modeling RTs in two-choice categorization paradigms rather than on analyzing perceptual dependencies. The purpose of the present paper is to develop a stochastic, dynamic GRT model for the full-factorial paradigm in order to focus directly on the perceptual dependencies at the heart of the original static theory. We return to a detailed consideration of the relationships between Ashby's model and our own in the discussion section at the end of the paper, where we also address the relationships between our model and those described by (e.g., Lamberts, 2000; Lamberts, Brockdorff, & Heit, 2003).

Our approach can be viewed as a generalization of most of the existing parameterized models of choice RTs, since it is based on a functional characterization of accumulation of evidence for a particular response until a particular bound is reached (hence the term "threshold-accrual-halting" for our model in the next section). Examples of such models include the accumulator models of Smith and Vickers (1988), the (unidimensional) random walk (e.g., Link & Heath, 1975) and the diffusion models of Ratcliff (1978), Ornstein–Uhlenbeck processes (e.g., Busemeyer & Townsend, 1993), and generalized diffusion processes such as those emanating from dynamic linear systems (e.g., Townsend & Wenger, 2004; Usher & McClelland, 2001).

The most novel aspects of the present theory are that theorems relating perceptual independence, perceptual separability, sampling independence, and marginal response invariance are established for functions of both RTs and accuracy. Of course these were only related through functions of accuracy in the original (static) theory. It will be seen straight away in that the key definitions underpinning the dynamic theory will be natural generalizations of those in the static theory. Thus, perceptual separability and perceptual independence must be defined in terms of the germane stochastic processes rather than static, joint distributions. Decisional separability possesses a quite logically similar meaning, but now in terms of the decision criteria on the separate activation channels. In addition, it will be seen that observable consequences are also analogous, but expanded in the stochastic situation, in the two theories.

Nonetheless, it is worth making an observation that might allay some confusion before beginning our technical exposition. When a theory is expanded to include additional dependent, observable variables, time in the present case, there is an apparent but not real paradox (a so-called antinomy). On the one hand, there is a true generalization of the theory since a greater spectrum of observable phenomena are covered. On the other hand and by the same token, models which obey the basic axioms may be more constrained since, in order to avoid falsification, they must predict data at a finer level of analysis.

Finally, the theorems will be concretized via simulations of a number of closely related stochastic linear dynamic models and a brief illustrative analysis of response time and accuracy data collected in two speech perception experiments (Silbert, 2010).

2. A threshold, accrual-halting parallel model

We now lay the groundwork for these new developments. We begin with the assumption that the information is processed in parallel with a conjunctive stopping rule.[1] At first glance, this may seem a rather restrictive assumption and there clearly is much work to be done investigating separability and independence in serial (and coactive) architectures, as well as with disjunctive stopping rules. However, we can already observe that in each new case there appear to be direct analogues. This follows because the key concepts, perceptual separability, decisional separability, and perceptual independence, appear to be more fundamental even than, say the architecture, although the details may differ somewhat. In any event, in the Discussion we will indicate the correspondents in the case of serial architectures.

For simplicity, we limit our analysis to two channels, each of which is composed of two subprocesses.[2] The channels are meant to represent processing of a particular dimension of a stimulus, while the subprocesses model evidence in favor of a particular decision on that dimension. A decision on a dimension is made whenever the first subprocess of that channel reaches a particular level of activation. The decision of the system is then the combination of the response on the two channels.

A summary of the notation is given in Table 1. We denote the state of the first channel at a particular time by $X(t)$ and assume that it can take any two-dimensional real value, $X(t) = (X_1(t), X_2(t)) \in \mathrm{R}^2$. Similarly, we denote the state of the second channel at time t by $Y(t) = (Y_1(t), Y_2(t)) \in \mathbb{R}^2$. We will also often use X and Y to refer to the channel whose state is represented by the $\{X(t); t \geq 0\}$ and $\{Y(t); t \geq 0\}$ processes. As an example, the first channel may represent a decision about color, with $X_1(t)$ representing accrual of evidence for the color red and $X_2(t)$ for the color purple. Then the second channel might represent a decision about shape, so $Y_1(t)$ could represent evidence for a square while $Y_2(t)$ represents evidence for a rectangle.

[1] Although Little, Nosofsky, and Denton (2011) provide compelling evidence that, in one particular categorization task, people will, when left to their own devices, tend to employ serial processing architectures with conjunctive stopping rules, it has been established since at least 1983 that processing architecture and stopping rule are logically independent (Townsend & Ashby, 1983). In the interest of conceptual cohesion and clarity, we leave most such work for future endeavors.

[2] More levels are possible per channel; we restrict the present work to two levels per channel for the sake of (relative) simplicity.

Table 1
List of symbols with brief definitions.

Symbol	Definition	Values
$X(t), Y(t)$	State of channel at time t	$\mathbb{R}^2$
$X_i(t), Y_i(t)$	State of subprocess i of channel at time t	$\mathbb{R}$
C_{x_i}, C_{y_i}	Decision criterion for subprocess i on channel	$\mathbb{R}^+$
s_x, s_y	Stimulus for a channel	$\{1, 2\}$
R_x, R_y	Response on a channel	$\{1, 2\}$
T	Completion time of the system	$\mathbb{R}^+$
T_x, T_y	Completion time of a channel	$\mathbb{R}^+$
T_b	Non-decision (base) time component of a response time	$\mathbb{R}^+$
$\mathrm{RT}(i, j)$	Response time for $R_x = i, R_y = j$	$\mathbb{R}^+$

The criterion on each subprocess is denoted by C with a subscript indicating the channel and subprocess, e.g., C_{x_2} indicates the criterion for the first channel's second subprocess. Completion times for each subprocess are given by the first passage time of the criterion of the process. Hence, the completion time of the first subprocess of the first channel, T_{x_1}, is given by $\inf\{t : X_1(t) \geq C_{x_1}\}$. The completion times on the other subprocess and channel are analogously defined. The channel completion time, T_x or T_y, is given by the completion time of earliest subprocesses to finish. Hence, the completion time of the first channel is $T_x = \min\left(T_{x_1}, T_{x_2}\right)$, and similarly for the other channel. The system has completed processing when both channels have completed. Since we assume the two channels operate in parallel, this implies that T is the larger of the channel completion times. Hence, $T_x = \min\left(T_{x_1}, T_{x_2}\right)$, $T_y = \min\left(T_{y_1}, T_{y_2}\right)$ and $T = \max\left(T_x, T_y\right)$. Observe that in the standard argot, processing on the two macro-channels (i.e., X, Y) is exhaustive whereas it is self-terminating (in fact, minimum time) on the subprocesses.

Now it is important to note that although our notation might suggest that we constrain each subprocess to parallel processing, this is not the case, our min time notation and express stochastic structures allow for arbitrary, for instance serial, architectures. With an assumption of serial processing within channel whichever activation, e.g., red or blue for a color dimension, is processed first will determine the dimensional choice, regardless of the stimulus values on that dimension. And degenerate special parallel models such as diffusion processes can, with appropriate substitutions, be written in terms of our general notation.[3]

The idea of subprocesses (stages, etc.) for within-dimension values is in accordance with all major processing models. For instance, a realization of the (Townsend & Ashby, 1983, Chapter 9) counter model in the current framework would allow for an X_1 and an X_2 counter. When the "1" value was presented then it would be expected that the X_1 counter rate would be higher than that for X_2. In the Ratcliff (1978) diffusion model, the drift rate, under the assumption of a presented "1" value, would favor motion toward the "1" barrier.[4]

The response ($R_x = i, R_y = j$), indicating level i on dimension X and level j on dimension Y, is determined by the channel completion events, though the time of the actual response – the RT – requires both these completion events as well as the base (or residual) time, $\mathrm{RT}(i, j) = T + T_b$. In some cases we will be interested in the marginal response distribution as well. The response time, without regard to the response is denoted $\mathrm{RT}(\cdot, \cdot)$ or simply RT. Likewise, if we are interested in the response times for a particular response on dimension X, marginalized across responses on Y we use $\mathrm{RT}(i, \cdot)$.

[3] Due to space considerations, we must omit the details of such special cases.

[4] The computational details of how, say Xi values are produced are again beyond our current purview, just as they are in almost all current parameterized models. One natural direction to pursue would be the use of dimensional (featural, etc.) filters to capture such mechanisms.

We use notation for responses that maintains a distinction between the information on each dimension, although the actual 1–1 stimulus–response assignments may or may not be stated explicitly in terms of the individual dimensions. For instance, suppose the stimuli are color and shape: red vs. purple and square vs. rectangle. The experimenter might request that the observer respond, for example, "red square", "purple square", "red rectangle" or "purple rectangle", or, alternatively and more simply, "1", "2", "3", or "4" to indicate which of the four stimuli which he/she believes was presented. We use the lower case s and to indicate that the stimulus is specified, we use the conditional operator.

At this point, we can state the distribution of the overall finishing time jointly with a certain response.

Definition 1 (*Joint Cumulative Distribution Function for Distinct Threshold Accrual-Halting Parallel Model*)**.** Here we give the joint distribution for response 1 on the X and Y dimensions, but the same form holds for any response with the obvious substitutions.

$$
\begin{aligned}
&P\{\mathrm{RT}(1, 1) \leq t | s_x = k, s_y = m\} \\
&\quad = P\left\{R_x = 1, R_y = 1, T + T_b \leq t | s_x = k, s_y = m\right\} \\
&\quad = P\left\{T_{x_1} < T_{x_2}, T_{y_1} < T_{y_2}, \max\left[T_x, T_y\right] \right. \\
&\qquad \left. +\ T_b \leq t | s_x = k, s_y = m\right\} \\
&\quad = P\Big\{\inf\{\tau : X_1(\tau) \geq C_{x_1}\} \leq \inf\{\tau : X_2(\tau) \geq C_{x_2}\} \\
&\qquad \cap \inf\{\tau : X_1(\tau) \geq C_{x_1}\} \leq t - T_b \\
&\qquad \cap \inf\{\tau : Y_1(\tau) \geq C_{y_1}\} \leq \inf\{\tau : Y_2(\tau) \geq C_{y_2}\} \\
&\qquad \cap \inf\{\tau : Y_1(\tau) \geq C_{y_1}\} \leq t - T_b | s_x = k, s_y = m\Big\}.
\end{aligned}
$$

This joint distribution may be simplified under the assumption of decisional separability, selective influence, perceptual separability or perceptual independence. The remainder of the theoretical section is dedicated to defining these terms for this model.

2.1. Decisional separability

In the static theory, *decisional separability* was a key notion and its satisfaction permitted strong conclusions to be drawn in both the non-parametric (Ashby & Townsend, 1986) as well as parametric (e.g., Kadlec & Townsend, 1992b; Thomas, 1999) instantiations. Decisional separability was defined as the assumption that a decision on one dimension was invariant over the levels of other dimensions.

The question therefore arises as to what corresponds to decisional separability in the present dynamic theory. The answer appears to be that the criteria on one dimension should not depend on the state or criteria of the other dimension.

Definition 2. Decisional separability holds on dimension X if and only if $P\{C_{x_i} \leq \gamma | Y(t), C_{y_1}, C_{y_2}, s_x, s_y\} = P\{C_{x_i} \leq \gamma | s_x, s_y\}$ for $i \in \{1, 2\}$ and all t and the distribution does not vary with changes to s_x or s_y. Similarly, decisional separability holds on dimension Y if and only if $P\{C_{y_i} \leq \gamma | X(t), C_{x_1}, C_{x_2}, s_x, s_y\} = P\{C_{y_i} = \gamma | s_x, s_y\}$ for $i \in \{1, 2\}$ and all t and the distribution does not vary with changes to s_x or s_y.

Returning to our color/shape example, decisional separability on the color dimension means that the probability that a certain amount of information is enough to respond "red" or "purple" does not depend on the amount of information about the shape, the decision thresholds for the shape, or which of the four stimuli were presented.

Although there are possibly some very interesting effects of failures of decisional separability on these models, we do not explore these effects in this paper. There is not a straightforward, adequate approach to modeling decisional interactions in the present framework. For example, one obvious approach may be to introduce correlation between C_{x_i} and C_{y_i}, however there is an equivalent model with deterministic bounds in which the correlation appears in the accumulation processes (cf. Dzhafarov, 1993) so this type of failure would be unidentifiable. Indeed, certain types of decisional separability failures are not identifiable even in static GRT (Silbert & Thomas, in press). In the theoretical section, we assume that decisional separability holds. In the simulation section, all of the models have fixed thresholds. Beyond establishing a definition, we have chosen to put aside this important topic to conserve space and allow more focus on the perceptual effects within these models.

2.2. Perceptual separability and marginal response invariance

Perceptual separability captures the idea in the static theory of GRT that the perceptual effect of one stimulus dimension does not depend on the level of the stimulus of another dimension. It is defined as an invariance of the marginal distribution on a dimension across levels of the other dimension. Failure of perceptual separability occurs in static GRT (i.e., with respect to responses accuracy) if changing the level of one stimulus dimension (e.g., color) affects the marginal perceptual distribution associated with a different stimulus dimension. In our example of shapes and colors, a failure of perceptual separability could mean, for example, that the perceptual effect of presenting a square depends on whether it was purple or red. In the dynamic framework under development here, if the X and Y channels are independent, failure of separability may be due, for example, to the dependence of X_1 on s_y. However, non-separability could also come about indirectly through dependence of the X channel on the Y channel.[5] In static GRT, if the bivariate distributions are Gaussian, then the marginal distributions are invariant across differing covariances. In general, stochastic dependence could be associated with non-separability of the marginal distributions, even in the static case. For an alternative approach to testing perceptual separability, see Dzhafarov (2002).

Here, we choose to follow the earlier track in our definition and maintain a distinction between separability and independence.

Definition 3. Perceptual separability of one channel at a particular stimulus level is defined as the invariance of the marginal processes of that channel over changes in the stimulus level of the other channel. Suppose $X(t, s_x = k, s_y = m)$ is the stochastic process corresponding to the activation in channel X at time t under stimulus conditions $s_x = k, s_y = m$. Then for perceptual separability to hold on channel X with stimulus $s_x = k$,

$$X(t, s_x = k, s_y = 1) \overset{d}{=} X(t, s_x = k, s_y = 2).$$

Similarly, for perceptual separability to hold on Channel Y with stimulus $s_y = m$,

$$Y(t, s_x = 1, s_y = m) \overset{d}{=} Y(t, s_x = 2, s_y = m).$$

We use $\overset{d}{=}$ to mean equal in distribution, by which we mean that the finite dimensional distributions of the two stochastic processes are equal (i.e., for all finite subsets of $[0, \infty]$ the distributions of the random variables are equal).

[5] This distinction is in line with the analogous notions in our stochastic theory of "direct vs. indirect non-selective influence" (Townsend & Ashby, 1983, Chapter 12).

Thus, perceptual separability of the color channel when the stimulus is red means that the cumulative distribution function (CDF) of the amount of time it takes a participant to accumulate a certain amount of information for "red" or "purple" when a red stimulus is presented, is the same whether the shape presented was a square or rectangle.

Note that perceptual separability may be asymmetric. X, or even just X_1, may be perceptually separable from Y while Y is not, in turn, perceptually separable from X, as can be seen in the specificity in the definition above.

Here, we would like to highlight the difference between perceptual separability and a related concept, direct separability. Conceptually, if direct separability holds, then the stimulus on one channel is not any part of the *input* to the other, but could influence the other through channel interactions. More formally, if direct separability holds for the first channel at level j,

$$P\{X_i(t)|Y(t), s_x = k, s_y = 1\}$$
$$= P\{X_i(t)|Y(t), s_x = k, s_y = 2\} \quad \text{for } i = 1, 2.$$

Whereas perceptual separability is a statement about *marginal* distributions, the definition of direct separability is about conditional distributions. Most importantly, direct separability does not imply perceptual separability. When direct separability holds, but the stimulus on one channel affects the state of the other channel due to interactions between the channels, perceptual separability fails. We refer to this as an indirect failure of separability. However, with the additional assumption of channel independence, direct separability is enough to imply perceptual separability.

Proposition 1. *If X and Y are independent and direct separability holds, then X and Y are perceptually separable at each level.*

We omit a formal proof, as Proposition 1 follows directly from the definitions.

Types of stochastic dependence brought about through the interaction of the channel processes (e.g., accumulating information or activation) would be expected to rarely, if ever, result in perceptual separability.

In static GRT, marginal response invariance is an observable statistic that is closely related to perceptual separability. In fact, marginal response invariance so cogently suggests *perceptual* invariance of a psychological dimension across levels of another that it earlier was used even for the theoretical notion itself, under the designation "across-stimulus invariance" (e.g., Townsend, Hu, & Ashby, 1981). The following simply recapitulates that original definition as stated in Ashby and Townsend (1986). It is straightforward to write the condition for response R_x given s_x, marginalized over R_y, and made to be the same for each of the two levels of s_y.

Definition 4. Marginal response invariance on a channel holds if and only if the marginal probability of a particular response on that dimension is invariant across the level of the other stimulus dimension. For marginal response invariance to hold on channel X,

$$P\{R_x = i|s_x = k, s_y = 1\} = P\{R_x = i|s_x = k, s_y = 2\}.$$

That is,

$$P\{R_x = i, R_y = 1|s_x = k, s_y = 1\}$$
$$+ P\{R_x = i, R_y = 2|s_x = k, s_y = 1\}$$
$$= P\{R_x = i, R_y = 1|s_x = k, s_y = 2\}$$
$$+ P\{R_x = i, R_y = 2|s_x = k, s_y = 2\}.$$

For marginal response invariance to hold on Channel Y,

$$P\{R_y = i|s_x = 1, s_y = m\} = P\{R_y = i|s_x = 2, s_y = m\}.$$

That is,

$$P\{R_x = 1, R_y = i|s_x = 1, s_y = m\} + P\{R_x = 2, R_y = i|s_x = 1, s_y = m\} = P\{R_x = 1, R_y = i|s_x = 2, s_y = m\} + P\{R_x = 2, R_y = i|s_x = 2, s_y = m\}.$$

For our color–shape example, marginal response invariance on the color channel means that the probability that a participant responds "red" is the same whether the shape presented was a square or a rectangle.

Now, given that we possess structure on processing times, we may ask if there is something that corresponds to marginal response invariance but also involves time. One such possibility is given in the next definition.

Definition 5. Timed marginal response invariance is defined by satisfaction of the condition for all $t > 0$,

$$P\{R_x = i, T \le t|s_x = k, s_y = 1\} = P\{R_x = i, T \le t|s_x = k, s_y = 2\}.$$

Equivalently,

$$P\{R_x = i, R_y = 1, T \le t|s_x = k, s_y = 1\} + P\{R_x = i, R_y = 2, T \le t|s_x = k, s_y = 1\} = P\{R_x = i, R_y = 1, T \le t|s_x = k, s_y = 2\} + P\{R_x = i, R_y = 2, T \le t|s_x = k, s_y = 2\}. \quad (1)$$

Thus, timed marginal response invariance means that the probability that a participant responds "red" at or before any particular time is the same whether the stimulus was a square or a rectangle. Like ordinary marginal response invariance, this equality is something that can be directly assessed using data.

Timed marginal response invariance assures us that the probability of the joint event that both the response on the X dimension is the ith level and the RT is not greater than time t, given arbitrary X stimulus level k, is invariant across levels of Y. Although it would be useful to have a comparable definition for an arbitrary X response with the X process finishing at time less than or equal to t, irrespective of when Y finishes, such events are unobservable in this design, as the system as a whole only finishes when *both* X and Y have finished processing.

Proposition 2. *Timed marginal response invariance implies ordinary marginal response invariance but not conversely.*

Proof. To see that timed marginal response invariance implies ordinary marginal response invariance, take the limit $t \to \infty$ of Eq. (1) in Definition 5. To see that the converse is not true in general, let t_0 be the median of T. Now consider the case in which

$$P\{R_x = 1|T \le t_0, s_x = 1, s_y = 1\} \ne P\{R_x = 1|T > t_0, s_x = 1, s_y = 1\}$$
$$P\{R_x = 1|T \le t_0, s_x = 1, s_y = 1\} = P\{R_x = 1|T > t_0, s_x = 1, s_y = 2\}$$
$$P\{R_x = 1|T \le t_0, s_x = 1, s_y = 2\} = P\{R_x = 1|T > t_0, s_x = 1, s_y = 1\}.$$

It is clear that timed marginal response invariance fails in this case. Ordinary marginal response invariance does still hold because,

$$\begin{aligned} P\{R_x = 1|s_x = 1, s_y = 1\} &= P\{R_x = 1|T \le t_0, s_x = 1, s_y = 1\}P\{T \le t_0\} \\ &\quad + P\{R_x = 1|T > t_0, s_x = 1, s_y = 1\}P\{T > t_0\} \\ &= P\{R_x = 1|T > t_0, s_x = 1, s_y = 2\}P\{T \le t_0\} \\ &\quad + P\{R_x = 1|T \le t_0, s_x = 1, s_y = 2\}P\{T > t_0\} \\ &= P\{R_x = 1|T > t_0, s_x = 1, s_y = 2\}P\{T > t_0\} \\ &\quad + P\{R_x = 1|T \le t_0, s_x = 1, s_y = 2\}P\{T \le t_0\} \\ &= P\{R_x = 1|s_x = 1, s_y = 2\}. \quad \square \end{aligned}$$

We now prove a theorem with sufficient conditions to ensure timed-marginal response invariance. In ordinary static GRT, perceptual and decisional separability are all that is required for MRI, but in this enlarged context, more subtlety appears. The response time is based on both the X and Y dimensions. Without an additional assumption, it is possible that, while the stimulus on the Y dimension does not affect the speed and response on the X dimension, the actual response time does depend on the Y stimulus. In the static GRT tests, this issue does not arise because one can isolate the response on X from the response on Y.

Definition 6. The speed invariance condition holds on Channel Y if $P\{T_y \le t|s_y = 1\} = P\{T_y \le t|s_y = 2\}$.

This requires that the marginal distribution on the completion time must stay unchanged across levels of the stimulus. Of course, the probability of a given response and how quickly that response is made should depend on the stimulus. This condition merely states that the amount of time it takes to make *any* response does not change with the stimulus. While this may seem like a strong constraint, the speed invariance condition is satisfied by a common assumption made in accumulator models, the assumption that the mean rate of information accumulation for the correct response is the same across stimuli within an experiment.

Proposition 3. *If perceptual and decisional separability hold and the speed invariance condition (Definition 6) holds on Channel Y, then timed marginal response invariance holds on channel X in accrual halting parallel models.*

Proof. Decisional separability implies that the criteria on one channel are not affected by a change in the input to the other channel,

$$P\{C_{x_1} = \gamma_1|s_x = k, s_y = 1\} = P\{C_{x_1} = \gamma_1|s_x = k, s_y = 2\}$$
$$P\{C_{x_2} = \gamma_2|s_x = k, s_y = 1\} = P\{C_{x_2} = \gamma_2|s_x = k, s_y = 2\}.$$

Suppose $t > 0$. Perceptual separability implies that the subprocesses on one channel are not affected by a change in the input to the other channel. So, along with decisional separability, this implies that,

$$P\{X_1(t) \ge C_{x_1}|s_x = k, s_y = 1\} = P\{X_1(t) \ge C_{x_1}|s_x = k, s_y = 2\}$$
$$P\{X_2(t) \ge C_{x_2}|s_x = k, s_y = 1\} = P\{X_2(t) \ge C_{x_2}|s_x = k, s_y = 2\}.$$

The distributions of first passage times are determined by the criteria and the subprocesses, hence,

$$\begin{aligned} &P\{T_{x_1} \le t, T_{x_1} < T_{x_2}|s_x = k, s_y = 1\} \\ &= P\Big\{\inf\{t' : X_1(t') \ge C_{x_1}\} \le t \cap \inf\{t' : X_1(t') \ge C_{x_1}\} \\ &< \inf\{t' : X_2(t') \ge C_{x_2}\}|s_x = k, s_y = 1\Big\} \\ &= P\Big\{\inf\{t' : X_1(t') \ge C_{x_1}\} \le t \cap \inf\{t' : X_1(t') \ge C_{x_1}\} \\ &< \inf\{t' : X_2(t') \ge C_{x_2}\}|s_x = k, s_y = 2\Big\} \\ &= P\{T_{x_1} \le t, T_{x_1} < T_{x_2}|s_x = k, s_y = 2\}. \end{aligned}$$

By the speed invariance condition on Y, $P\{T_y \le t|s_y = 1\} = P\{T_y \le t|s_y = 2\}$. Therefore,

$$\begin{aligned} &P\{R_x = 1, T \le t|s_x = k, s_y = 1\} \\ &= P\{T_{x_1} \le t, T_{x_1} < T_{x_2}, T_y \le t|s_x = k, s_y = 1\} \\ &= P\{T_{x_1} \le t, T_{x_1} < T_{x_2}, T_y \le t|s_x = k, s_y = 2\} \\ &= P\{R_x = 1, T \le t|s_x = k, s_y = 2\}. \quad \square \end{aligned}$$

Despite the additional assumption needed in Proposition 3 and the model based definition of perceptual separability (Definition 3), perceptual separability and decisional separability are sufficient to imply static marginal response invariance, maintaining consistency with earlier definitions.

Proposition 4. *Perceptual separability and decisional separability imply marginal response invariance in accrual halting parallel models.*

Proof. Decisional separability implies that the criteria on one channel are not affected by a change in the input to the other channel,

$$P\{C_{x_1} = \gamma_1 | s_x = k, s_y = 1\} = P\{C_{x_1} = \gamma_1 | s_x = k, s_y = 2\}$$
$$P\{C_{x_2} = \gamma_2 | s_x = k, s_y = 1\} = P\{C_{x_2} = \gamma_2 | s_x = k, s_y = 2\}.$$

Perceptual separability implies that the subprocesses on one channel are not affected by a change in the input to the other channel. So, along with decisional separability, this implies that,

$$P\{X_1(t) \geq C_{x_1} | s_x = k, s_y = 1\} = P\{X_1(t) \geq C_{x_1} | s_x = k, s_y = 2\}$$
$$P\{X_2(t) \geq C_{x_2} | s_x = k, s_y = 1\} = P\{X_2(t) \geq C_{x_2} | s_x = k, s_y = 2\}.$$

The distributions of first passage times are determined by the criteria and the subprocesses, so,

$$\begin{aligned} &P\{T_{x_1} < T_{x_2} | s_x = k, s_y = 1\} \\ &= P\left\{\inf\{t : X_1(t) \geq C_{x_1}\} \right. \\ &\left. < \inf\{t : X_2(t) \geq C_{x_2}\} | s_x = k, s_y = 1\right\} \\ &= P\left\{\inf\{t : X_1(t) \geq C_{x_1}\} \right. \\ &\left. < \inf\{t : X_2(t) \geq C_{x_2}\} | s_x = k, s_y = 2\right\} \\ &= P\{T_{x_1} < T_{x_2} | s_x = k, s_y = 2\}. \end{aligned}$$

Thus, because the response of a channel is determined by the first passage time of the subprocesses,

$$\begin{aligned} P\{R_x = 1 | s_x = k, s_y = 1\} &= P\{T_{x_1} < T_{x_2} | s_x = k, s_y = 1\} \\ &= P\{T_{x_1} < T_{x_2} | s_x = k, s_y = 2\} \\ &= P\{R_x = 1 | s_x = k, s_y = 2\}. \end{aligned}$$

Therefore, together, decisional separability and perceptual separability imply that the response on a channel is not affected by a change in the input to the other channel and hence marginal response invariance will hold. □

2.3. Perceptual independence and report (sampling) independence

Perceptual independence in static GRT is defined as stochastic independence of the joint distribution of observations on the different dimensions. Note that this condition is typically not directly observable. It is natural to extend this concept to the present framework via the following definition.

Definition 7. Two channels are said to be perceptually independent if $\{X(t); t \geq 0\}$ and $\{Y(t); t \geq 0\}$ are independent.[6]

The shape and color channels in our example are perceptually independent if the amount of information a participant has accumulated for a response on the shape dimension has no bearing on the amount of information accumulated on the color dimensions and vice versa.

[6] This means that the finite dimensional distributions of X and Y are independent, so for any $\tau \in [0, \infty)$ with cardinality of $n < \infty$ and $A, B \in \mathbb{R}^n$, then $P(\{X(t)\}_{t\in\tau} \in A \text{ and } \{Y(t)\}_{t\in\tau} \in B) = P(\{X(t)\}_{t\in\tau} \in A) \times P(\{Y(t)\}_{t\in\tau} \in B)$.

Referring to an observable correlate of perceptual independence, the term "sampling independence" was first used in the context of tests of axioms about feature extraction from letters and letter-like stimuli (Townsend et al., 1981; Townsend, Hu, & Evans, 1984; Townsend, Hu, & Kadlec, 1988). While it would be desirable to maintain the terminology from earlier papers on GRT, the term sampling independence connotes an independence within the system. However, like marginal response invariance, sampling independence is observable. Because of its observability, we shall refer to sampling independence here as "report independence".

In both discrete, feature-based models and continuous, multidimensional models, report independence has an important role to play. Under appropriate assumptions, such as the presence of decisional separability in static GRT, report independence and perceptual independence are equivalent. A particular interesting model satisfying report independence occurs when a finite set of discrete features are extracted independently and the observer reports exactly what he/she extracts (e.g., Schulze, Baurichter, Gerling, & Grobe, 1977; Townsend et al., 1981; Wandmacher, 1976). The definition of report independence follows.

Definition 8. We say that report independence holds for a particular stimulus–response combination if the probability of that response is equal to the product of the marginal probability of each component of the response. Formally, report independence holds for $R_x = 1, R_y = 1$ with stimulus $s_x = k, s_y = m$ if,

$$\begin{aligned} &P\{R_x = 1, R_y = 1 | s_x = k, s_y = m\} \\ &= P\{R_x = 1 | s_x = k, s_y = m\} \times P\{R_y = 1 | s_x = k, s_y = m\}. \end{aligned}$$

Equivalently, in terms of observable quantities,

$$\begin{aligned} &P\{R_x = 1, R_y = 1 | s_x = k, s_y = m\} \\ &= \left[P\{R_x = 1, R_y = 1 | s_x = k, s_y = m\} \right. \\ &\left. + P\{R_x = 1, R_y = 2 | s_x = k, s_y = m\}\right] \\ &\times \left[P\{R_x = 1, R_y = 1 | s_x = k, s_y = m\} \right. \\ &\left. + P\{R_x = 2, R_y = 1 | s_x = k, s_y = m\}\right]. \end{aligned}$$

Report independences for the other stimulus–response combinations are analogous.

In color and shape terms, report independence means that the probability of a participant responding "red" and "square" to a particular stimulus is equal to the marginal probability that he responds "red" to that stimulus (i.e., either "red-square" or "red-rectangle") multiplied by the marginal probability that he responds "square" to that stimulus.

Similar to timed marginal response invariance, we would like to extend the concept of report independence to the response time domain. We use the same approach here, requiring that the static version hold for all t. Framing the definition in terms of the CDFs of response times leads to one important difference between the static and dynamic version. We cannot simply require that timed report independence holds if the probability of a response before some time is equal to the product of the probabilities that each channel had the indicated response before that time, i.e.,

$$\begin{aligned} &P\{\mathrm{RT}(i, j) \leq t | s_x = k, s_y = m\} \\ &\neq P\{\mathrm{RT}(i, \cdot) \leq t | s_x = k, s_y = m\} \\ &\quad \times P\{\mathrm{RT}(\cdot, j) \leq t | s_x = k, s_y = m\}. \end{aligned} \tag{2}$$

This is because each of the marginal terms on the right hand side include the probability that the overall response time was less than t. One possibility would be to define timed report independence in terms of the channel completion times, however this would lead to an unobservable condition. Instead, we use the response probabilities conditioned on response times.

Definition 9. We say that timed report independence holds for a particular stimulus–response combination if the conditional probability of the response, conditioned on the response having been made by t, is equal to the product of the marginal probability of each component of the response, each conditioned on the response time being less than t, for all $t > 0$. Formally, report independence holds for $R_x = i, R_y = j$ with stimulus $s_x = k, s_y = m$ if for all $t > 0$,

$$P\{R_x = i, R_y = j | \mathrm{RT}(\cdot,\cdot) \le t, s_x = k, s_y = m\}$$
$$= P\{R_x = i | \mathrm{RT}(\cdot,\cdot) \le t, s_x = k, s_y = m\}$$
$$\times P\{R_y = j | \mathrm{RT}(\cdot,\cdot) \le t, s_x = k, s_y = m\}.$$

Equivalently,

$$P\{\mathrm{RT}(i,j) \le t | s_x = k, s_y = m\}$$
$$\times P\{\mathrm{RT}(\cdot,\cdot) \le t | s_x = k, s_y = m\}$$
$$= P\{\mathrm{RT}(i,\cdot) \le t | s_x = k, s_y = m\}$$
$$\times P\{\mathrm{RT}(\cdot,j) \le t | s_x = k, s_y = m\}.$$

For shapes and colors, timed report independence means that the probability that a participant has responded "red" and "square" to a stimulus given that she has responded by a particular time is equal to the product of the probability that she responded "red" and the probability that she responded "square", each on the condition that she responded by that same time. The second formulation in the definition gives a way that this property can be tested with data. The joint probability of a "red-square" response and the response occurring by a particular time multiplied by the probability that any response was made by that time is the probability that any "red" response was made by that time multiplied by the probability that any "square" response was made by that time.

Much like the connection between perceptual separability and marginal response invariance, there is also a connection between perceptual independence and timed report independence.

Proposition 5. *Perceptual independence and decisional separability imply timed report independence in accrual halting parallel models.*

Proof. For convenience of notation, we assume that the density of the residual times, f_b, exists, but it is not necessary for the theorem to hold.

$$P\{\mathrm{RT}(i,j) \le t | s_x = k, s_y = m\}$$
$$= P\{R_x = i, R_y = j | \mathrm{RT} \le t, s_x = k, s_y = m\}$$
$$= P\{R_x = i, R_y = j | T_x \le t - T_b,$$
$$T_y \le t - T_b, s_x = k, s_y = m\}$$
$$\times P\{T_x \le t - T_b, T_y \le t - T_b | s_x = k, s_y = m\}$$
$$= \int_0^\infty P\{R_x = i, R_y = j | T_x \le t - t_b,$$
$$T_y \le t - t_b, s_x = k, s_y = m\}$$
$$\times P\{T \le t - t_b\} f_b(t_b)\, dt_b$$
$$= \int_0^\infty P\{R_x = i | T_x \le t - t_b, T_y \le t - t_b, s_x = k, s_y = m\}$$
$$\times P\{R_y = j | T_x \le t - t_b, T_y \le t - t_b, s_x = k, s_y = m\}$$
$$\times P\{T \le t - t_b\} f_b(t_b)\, dt_b$$
$$= P\{R_x = i | T_x \le t - T_b, T_y \le t - T_b, s_x = k, s_y = m\}$$
$$\times P\{R_y = j | T_x \le t - T_b, T_y \le t - T_b, s_x = k, s_y = m\}$$
$$\times P\{\mathrm{RT} \le t\}$$
$$= P\{\mathrm{RT}(i,\cdot) | s_x = k, s_y = m\}$$
$$\times P\{R_y = j | T_x \le t - T_b, T_y \le t - T_b, s_x = k, s_y = m\}$$
$$= P\{\mathrm{RT}(i,\cdot) | s_x = k, s_y = m\}$$
$$\times P\{\mathrm{RT}(\cdot,j) | s_x = k, s_y = m\} / P\{\mathrm{RT} \le t\}$$
$$\text{for } P\{\mathrm{RT} \le t\} > 0.$$

Hence,

$$P\{\mathrm{RT}(i,j) \le t | s_x = k, s_y = m\} P\{\mathrm{RT} \le t | s_x = k, s_y = m\}$$
$$= P\{\mathrm{RT}(i,\cdot) \le t | s_x = k, s_y = m\}$$
$$\times P\{\mathrm{RT}(\cdot,j) \le t | s_x = k, s_y = m\}. \quad \square$$

Thus, if we find violations of timed report independence, either decisional separability or perceptual independence must have failed.

Like timed marginal response invariance, timed report invariance is more general than standard report invariance in these models. We do not include a proof as it is essentially the same as the proof of Proposition 2.

Proposition 6. *Timed report independence implies ordinary report independence but not conversely.*

In this case, we did not need an additional constraint to extend the report independence test to include response times. Despite the model based definition of perceptual independence (Definition 7), perceptual independence and decisional separability are sufficient to imply static report independence by Propositions 5 and 6. Hence, the new approach maintains consistency with earlier definitions.

2.4. Summary of theoretical results

We have described general theoretical machinery which extends static General Recognition Theory to the time domain. More specifically, the machinery represents a generalization of earlier efforts to develop stochastic, dynamic GRT models (e.g., Ashby, 1989, 2000), with a focus on the relationships between psychological dimensions and the implications of these relationships for observable statistics.

In the original static GRT, decisional separability conjoined with perceptual separability imply marginal response invariance. In the present Proposition 3, we also require that the marginal invariance condition be satisfied in order to imply timed marginal response invariance. With reference to timed report independence, we have a complete analogue to static report independence in the static domain, since decisional separability plus perceptual independence logically entail timed report independence.

3. Simulations via linear stochastic dynamic systems

The theorems given above establish a set of sufficient conditions for timed marginal response invariance and timed report independence. In order to more fully explore the ways in which the underlying relationships between processing channels influence timed marginal response invariance and timed report independence, response times were simulated with various parameterizations of a linear stochastic dynamic systems model.

We explore two types of noise in this model, early noise which is integrated along with the signal as part of the information accumulation process and a late noise which is added on to the accumulated evidence before the comparison to the threshold. The first noise source is meant to represent uncertainty in the input, including the uncertainty due to transduction of the physical stimulus. We model this noise with a white noise process and thus, depending on whether or not there is a feedback parameter in the model, the accumulated information is modeled as either a Brownian motion process with drift or an Ornstein–Uhlenbeck process. Following Ashby, we refer to the second noise source as the perceptual noise. We model this noise as a Brownian motion process (with no drift).[7]

One of the goals of this section is to explore the individual effect of each type of noise on channel dependencies, and particularly on timed marginal response invariance and timed report independence, the empirical statistics related to perceptual separability and perceptual independence. As such, we did not simulate models with both noise sources present (although a model with both types of noise would be quite reasonable and perhaps even a better model of true cognitive processes).

Figs. 1, 2 and 4 through 7 depict variations on the model and the general model can be described formally by the following equations,

$$d\mathbf{w}(t) = \mathbf{A}\mathbf{w}(t)\,dt + \mathbf{u}(t)\,dt + d\boldsymbol{\eta}_a(t) \quad (3)$$

$$\mathbf{v}(t) = \mathbf{w}(t) + \boldsymbol{\eta}_b(t). \quad (4)$$

In these equations, $\mathbf{u}$ is the vector of inputs to the system, η_a is the input noise and η_b is the perceptual noise. The elements of the matrix $\mathbf{A}$ govern interactions between the channels. If $\mathbf{A}$ is diagonal and the noise sources are independent, the channels are independent and separable.[8] Various sorts of interactions between the channels can be implemented via a variety of configurations of non-zero off-diagonal elements in $\mathbf{A}$; a number of these are described below in detail, as are a number of models with diagonal $\mathbf{A}$ and correlated noise added. As the definition of the model is otherwise symmetric, the speed invariance condition (Definition 6) will hold as long as $\mathbf{A}$ is symmetric.

Simulations were run in a four channel linear stochastic system, the channels corresponding to each of two levels on two dimensions: X_1, X_2, Y_1 and Y_2. As described above, though, the model can be readily extended to an arbitrary number of dimensions and an arbitrary number of levels on each dimension. Because the model has four channels, and any channel may, in theory, interact with any other, it is difficult to depict the model clearly with all possible interactions. A schematic of each of the models discussed is included along with their associated $\mathbf{A}$ matrices, noise insertion points, and noise covariance structures.

Fig. 1 shows a schematic representation of a fully independent and separable model. Consider the topmost channel in this schematic representation. Here u_{x_i} represents the incoming information in favor of response i on channel X. We remain agnostic on the exact form of the translation from physical stimulus to u other than that u is a non-decreasing function of the objective evidence available for its corresponding response. Although it is not necessary, we treat u_{X_i} and u_{X_j} as unrelated, then model any connections via interactions from the $\mathbf{A}$ matrix. $d\eta_{x_{1a}}$ represents (white) input noise, $\eta_{x_{1b}}$ represents (Brownian) perceptual noise, W_{x_1} represents perceptual activation, V_{x_1} the activation after the addition of perceptual noise (i.e., $V_{x_1} = W_{x_1}$ if $\eta_{x_{1b}} = 0$). The other three channels have analogous corresponding elements. The a parameters in the independent, separable model are the diagonal elements of the $\mathbf{A}$ matrix from Eq. (3), multipliers

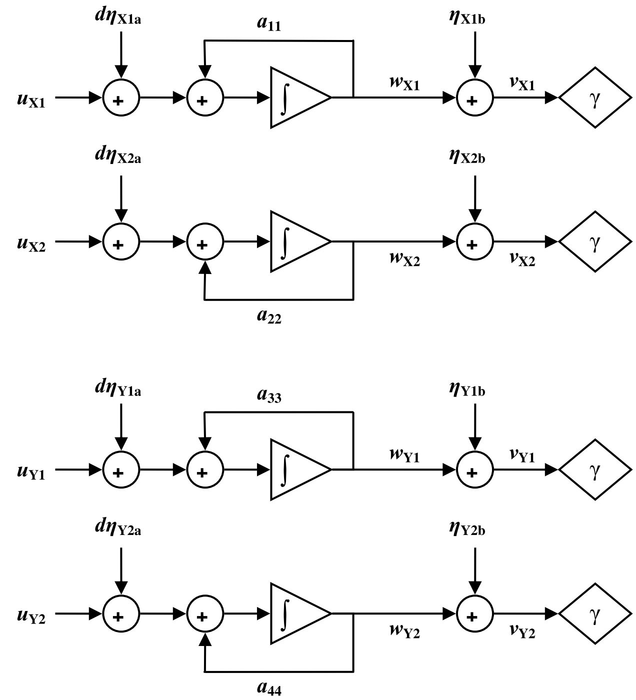

Fig. 1. Schematic representation of independent and separable four channel linear stochastic dynamic system used to simulate response time data. u_{Xi} represents incoming evidence in favor of response i on channel X (and likewise for Y). a_{ij} represents the entry in the $\mathbf{A}$ matrix that governs within sub-process feedback interactions among the processes. $d\eta_a$ and η_b are the input and perceptual noise respectively.

that provide stability via small magnitude negative feedback within each channel. The triangles represent integrators, and the diamonds represent activation (decision) criteria. In the model depicted in Fig. 1, each channel processes its own input (if present) and (input or perceptual) noise, with no influence from any of the other channels. The model is independent, so timed report independence will hold (Proposition 3); the speed invariance condition holds and the model is separable, so timed marginal response invariance will hold (Proposition 5).

By way of contrast, Fig. 2 depicts a model with bidirectional interactions, on one hand, between X_1 and X_2 and, on the other hand, between Y_1 and Y_2. In this model, a_{12} and a_{21} govern the interactions between the X channels, and a_{34} and a_{43} govern the interactions between the Y channels. Suppose, for example, that $a_{12} < 0$ and $a_{21} < 0$. In this case, input $s_x = 1$ would increase the activation in X_1 and decrease the activation in X_2. Similarly, input $s_x = 2$ would increase the activation in X_2 and decrease the activation in X_1. Both would serve to reduce the number of errors on the X dimension. On the other hand, if $a_{12} > 0$ and $a_{21} > 0$, increased activation in either X channel would correspond to increased activation in the other X channel, and the number of errors on the X dimension would increase. Analogous logic holds for the Y channels and the a_{34} and a_{43} parameters. Because these interactions are within channel, rather than X affecting Y or vice versa, this model is still perceptually separable. As long as the noise sources are independent, the model will also be perceptually independent. Hence, timed report independence will hold (Proposition 5). As long as the $\mathbf{A}$ matrix is balanced in such a manner that the speed-invariance condition holds, marginal response invariance will also hold (Proposition 3).

Fig. 4 depicts a model with unidirectional interactions between X_1 and Y_1, via the a_{31} parameters, and between X_2 and Y_2, via the a_{42} parameter. In this model, the activation in each X channel influences the activation in each corresponding Y channel, but not

7 One might consider a white noise process as an alternative to a Brownian motion process for perceptual noise, so as to match the input noise. Because the input noise is integrated before the threshold comparison stage, a Brownian motion process has more comparable effects on the completion time distributions. Adding a white noise process at this stage in the model leads to completion time distributions that are unlike empirically observed response time distributions: To have a noticeable effect on the completion time distribution, large variance values are needed. However, these large variances drown out the effect of the stimulus and lead to the majority of sample paths reaching the threshold extremely quickly.

8 Ashby (1989) shows, in the analogous discrete model, that independence also occurs if $\Sigma^{-\frac{1}{2}}(\mathbf{I} - \mathbf{A})$ is orthogonal, where Σ is the (diagonal) covariance matrix of the noise added to the model and $\mathbf{I}$ is the identity matrix.

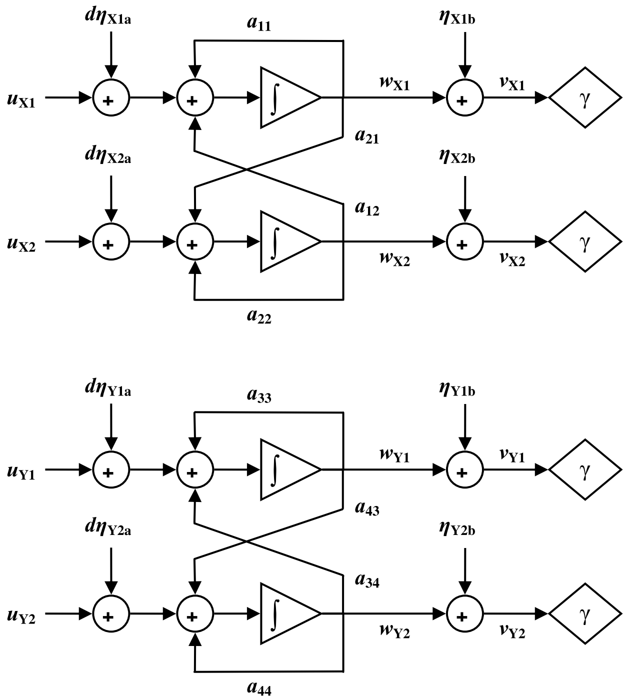

Fig. 2. Schematic representation of four channel linear stochastic dynamic system with bidirectional interactions between X_1 and X_2 and between Y_1 and Y_2.

vice versa. Suppose, for example, that $a_{31} > 0$ and $a_{42} > 0$. In this case, $s_x = 1$ would increase activation in both X_1 and Y_1, increasing the accuracy of responses to the $s_x = 1, s_y = 1$ stimulus and decreasing the accuracy of responses to the $s_x = 1, s_y = 2$ stimulus (due to an increase in erroneous $R_x = 1, R_y = 1$ responses). Similarly, input $s_x = 2$ would increase activation in both X_2 and Y_2, increasing the accuracy of responses to the $s_x = 2, s_y = 2$ stimulus, and decreasing the accuracy of responses to the $s_x = 2, s_y = 1$ stimulus (due to an increase in erroneous $R_x = 2, R_y = 2$ responses). A similar model with the Y channels influencing the X channels could be constructed via non-zero a_{13} and a_{24} parameters. Thus, in this model, perceptual separability fails, so we would not expect timed marginal response invariance to hold. If $d\eta_{x_a} > 0$, then the information fed from X to Y will also induce perceptual dependencies and we would not expect time report independence to hold.

The model used here is closely related to the model developed by Ashby (1989), though, as mentioned in the Introduction, there are some key differences. It is worth being explicit about three important differences between Ashby's linear stochastic GRT model and the present model. First, response-selection is implemented differently in the two models. Ashby (1989) addresses a number of different decision strategies, modeling the RTs and accuracy of two-choice decisions made under speed stress via multidimensional random walks. Extending the discrete time case to continuous time, Ashby (2000) also shows how multidimensional diffusion processes in two-choice classification can be reduced to a unidimensional diffusion with absorbing boundaries. Ashby (1989) also models identification-confusion probabilities (i.e., accuracy but not RTs) with larger sets of factorially constructed and tachistoscopically presented stimuli as a function of the relative activation levels in separate channels. By way of contrast, we model RTs and accuracy simultaneously in four-choice responses to factorially constructed stimuli, employing (equal, but distinct, and so in theory potentially unequal) decision criteria on each channel's activation. In addition to modeling distinct task types, the race-model is architecturally quite different from the random walk and diffusion models developed by Ashby. An additional, less significant difference, is that Ashby models activation in each channel as a power function of the input, whereas we model channel activation as the integral of the input. Next, whereas our simulation model is a discrete approximation to the continuous time model described above, the model described in Ashby (1989) is an explicitly discrete time model. Finally, the activation in our model may be negative as well as positive, though only a positive-valued absorbing barrier may stop the process on any given channel (cf. Usher & McClelland, 2001 who use a reflecting barrier at 0 to restrict activation to positive values).

In the four-channel model, there are 12 off-diagonal elements in the **A** matrix, which gives us an unwieldy 4096 distinct parameterizations implementing various kinds of interactions between the channels. Allowing for variation in the magnitude of the interactions, there is an infinite number of possible models. Obviously, a full exploration of this model space is both unrealistic and largely uninteresting. Only a small subset of these are of interest. These are the models exhibiting failures of perceptual independence and separability with direct analogs in static GRT. These will be described in detail below.

3.1. Simulation methods

Simulations were run with 14 models exhibiting various types of interactions. Two models exhibiting perceptual independence and perceptual separability were used to set a baseline against which to compare various models exhibiting failure of perceptual independence, failure of perceptual separability, or both. Failure of perceptual separability was implemented in two ways: by causing salience (i.e., perceptual distance) to vary on one dimension as a function of the other or by inducing so-called "mean-shift" integrality. This is a failure of separability in which the mean perceptual effect on one dimension is shifted as you go from one level to another on the other dimension. We will discuss mean-shift failures of separability (i.e., integrality) in more detail when the specific model is presented.

Asymmetric and symmetric salience-induced and pseudo mean-shift failures of separability were modeled. In asymmetric simulations, separability failed on only one dimension, whereas in symmetric simulations, separability failed on both dimensions simultaneously. For all but two of the methods for inducing failure of independence or separability, two models were run: one with input-noise only (i.e., $\boldsymbol{\eta}_a$ was non-zero and $\boldsymbol{\eta}_b$ was zero), and one with perceptual noise only (i.e., $\boldsymbol{\eta}_a$ was zero and $\boldsymbol{\eta}_b$ was non-zero). Each model is described in detail prior to presentation of the results.

For every model, 100 simulations were run. Each simulation consisted of 250 trials for each stimulus produced by a factorial combination of the two levels on each of the two channels (i.e., $(s_x = 1, s_y = 1)$, $(s_x = 1, s_y = 2)$, $(s_x = 2, s_y = 1)$, and $(s_x = 2, s_y = 2)$). Input was always present in exactly one X channel and one Y channel; there were no noise-only stimuli.

There is a large number of potentially free parameters in addition to those in the **A** matrix, many of which have complementary effects. Thus, a number of parameters were held constant. First, the time-step used in the simulations and the magnitude of the input were both set at unity. The duration of the input was set to 2000 time steps, and the within-channel feedback parameters (i.e., the diagonal of the **A** matrix in Eq. (3)) were set at −0.005. The variance of (each element of) $\boldsymbol{\eta}_a$ and $\boldsymbol{\eta}_b$ was set at 450. For simulations with input noise (non-zero $\boldsymbol{\eta}_a$), the decision criteria for each channel was set at 175, and for simulations with perceptual noise (non-zero $\boldsymbol{\eta}_b$), the decision criteria for each channel was set at 400. The noise variance and decision criteria values were arrived at after extensive trial and

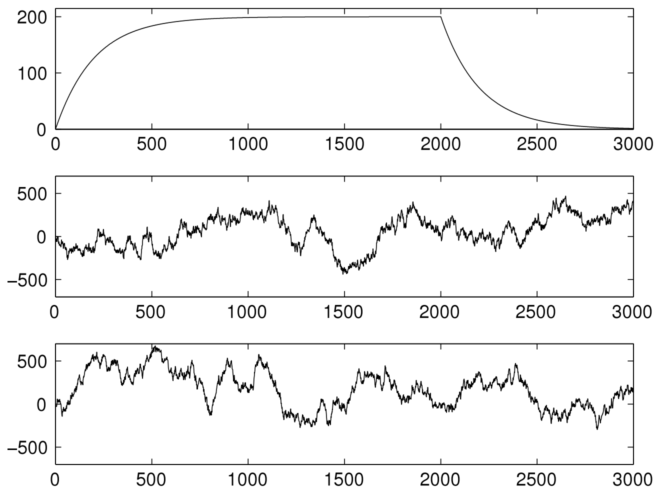

Fig. 3. Channel activation with no noise (top panel), with input noise only (η_a; middle panel), and with perceptual noise only (η_b; bottom panel).

error in an effort to produce reasonable identification accuracy levels in the simulated data. We added base times to the simulated response times, which were uniformly distributed on the interval [75, 225].

Fig. 3 shows, for a single trial, the activation of a channel with non-zero input and no noise (top panel), with input noise (middle panel), and with perceptual noise (bottom panel). The fixed parameter values cause the deterministic system to increase and approach 200 asymptotically, until the end of the input (time step 2000), at which point it decays exponentially. Note that, due to the relatively high variance of added noise, the behavior of the deterministic system is difficult to see clearly when noise is added to the system. Note, too, that the range of the y-axis in the bottom two panels is different than in the top panel of the figure.

For any given model, the statistics of interest characterize relationships between the joint distributions of response value and response time, which we occasionally will refer to as "defective" CDFs. The number of stimuli and their dimensional structure determines the number of comparisons of interest; again we restrict ourselves to two levels on each of two dimensions, though we see no reason, in principle, that the results presented here could not scale up to a larger number of levels on each of a larger number of dimensions.

We check for timed marginal response invariance (Definition 5) by comparing defective CDFs for correct responses to each level on each dimension across levels of the other dimension. This amounts to jointly testing whether either the response probabilities are different or the response time distributions are different. For example, for the first level on the X dimension, we compare $P\{R_x = 1, T \leq t|s_x = 1, s_y = 1\}$ to $P\{R_x = 1, T \leq t|s_x = 1, s_y = 2\}$, the defective CDFs for correct responses to the first level on the X dimension across levels of the Y dimension. We carry out analogous comparisons for each level on each dimension.

We tested these differences for significance using an adaptation of the Kolmogorov–Smirnov test. For each simulation we calculated the maximum absolute difference of the distribution, scaled by $(1/m + 1/n)^{-1/2}$, where m and n are the number of responses underlying the two defective CDFs under consideration, respectively. We then compare this test statistic to the Kolmogorov distribution.

Timed report independence results are similar, though the comparisons are between defective CDFs for correct responses on each dimension simultaneously, multiplied by the corresponding CDF for any response, and the product of marginal defective CDFs for correct responses on each component dimension (see Proposition 5 above):

$$\begin{aligned}&\left[P\{R_x = 1R_y = 1, \mathrm{RT} \leq t|s_x = 1, s_y = 1\}\right.\\&\quad\left.\times\ P\{RT \leq t|s_x = 1, s_y = 1\}\right]\\&\quad - \left[P\{R_x = 1, \mathrm{RT} \leq t|s_x = 1, s_y = 1\}\right.\\&\quad\left.\times\ P\{R_y = 1, \mathrm{RT} \leq t|s_x = 1, s_y = 1\}\right]. \qquad (5)\end{aligned}$$

To test for statistical significance, we used the input noise, independent, separable model to simulate a sampling distribution for the null-hypothesis.[9] For the simulations of each of the other models, we compared the value of the timed report independence test to the critical values derived from that simulated distribution.

We restrict ourselves to carrying out these comparisons only for correct responses, both to conserve space and mitigate the negative effects of the considerably smaller number of data points for incorrect responses. As with the timed marginal response invariance methods described above, the point here is to illustrate how timed report analysis is analyzed. The timed report independence results for all 14 models are presented in detail below.

Table 2
Percentage of simulations with significant results for independent separable model with input noise (η_a) or perceptual noise (η_b). Both perceptual separability and independence should hold for this model, so neither test should be significant. Results from the timed report independence test of the η_a model were used to estimate the sampling distribution under the null hypothesis that timed report independence holds, so they were not included here.

	Constant	Timed		Static	
		η_a	η_b	η_a	η_b
MRI	$R_x = 1; s_x = 1$	0	0	0.01	0
	$R_x = 2; s_x = 2$	0.02	0	0.03	0.02
	$R_y = 1; s_y = 1$	0	0	0.04	0.01
	$R_y = 2; s_y = 2$	0.01	0	0.05	0
RI	$s_x = 1; s_y = 1$	NA	0.11	0.02	0.07
	$s_x = 1; s_y = 2$	NA	0.04	0.07	0.04
	$s_x = 2; s_y = 1$	NA	0.04	0.07	0.06
	$s_x = 2; s_y = 2$	NA	0.09	0.03	0.07

3.2. Independence and separability in a baseline system

We establish a baseline set of results by running simulations with fully independent and separable models (depicted in Fig. 1), one with input noise and one with perceptual noise. These results serve as a point of comparison for the 12 models in which independence and/or separability fail. Following the description given above, the fully independent and separable models have no cross-talk between channels (i.e., **A** is diagonal). Hence, because the parameters governing the input, feedback, and noise do not vary as a function of level or dimension, and based on Propositions 3 and 5, timed marginal response invariance and timed report independence should hold. The simulations serve to confirm these results (on average) as well as provide a sense of how much variability there is across simulations.

Results for these models are summarized in Table 2. The first four rows correspond to tests of timed marginal response invariance for the correct marginal response to each level of the stimulus on each dimension. For each round of the simulation, the KS test described above was applied with $\alpha = .05$. The

[9] It would also have been reasonable to use the perceptual noise model to generate the null-hypothesis distribution. Given that the feedback parameters (a_{ii}) were small relative to the input/threshold ratio, the distribution would not have been very different.

number in the last column is the proportion of times the observed value was statistically significant. The 'constant' column denotes the responses and stimulus properties relevant to a given test as, e.g., $R_x = i$; $s_x = k$, indicating that the results report the proportion of statistically significant differences between $P\{R_x = i|s_x = k, s_y = 1\}$ and $P\{R_x = i|s_x = k, s_y = 2\}$, and analogously for responses and stimulus properties on the Y dimension.

The lower four rows correspond to tests of timed report independence for each level of the stimulus on each dimension. As described above, the input noise, independent, separable model was used to simulate a sampling distribution for the test of time report independence. Hence, the results for this test are necessarily 0.05. The results of the timed report independence test for the perceptual noise model give us an idea of the range of possible type-II error rates for this test. For these tests, $s_x = i$; $s_y = j$ in the 'constant' column indicates that the results report the proportion of statistically significant differences between $P\{\mathrm{RT}(i, j) \leq t|s_x = i, s_y = j\}P\{\mathrm{RT}(\cdot, \cdot) \leq t|s_x = i, s_y = j\}$ and $P\{\mathrm{RT}(i, \cdot) \leq t|s_x = i, s_y = j\}P\{\mathrm{RT}(\cdot, j) \leq t|s_x = i, s_y = j\}$.

These results are as expected. Both perceptual separability and perceptual independence hold for this model, and the corresponding test statistics were significant rarely, roughly at the frequency expected for tests with $\alpha = .05$. It is worth noting, that the implicational relationship between timed and static marginal response invariance stated in Proposition 2 appears to hold for the baseline model with either input or perceptual noise. On the other hand, we see two violations of the analogous implication between timed and static report independence, though both are of small magnitude. Differences in the power of the static and timed tests may account for such discrepancies with the theoretically expected outcomes.

3.3. Failure of separability due to changes in salience

In static General Recognition Theory, separability may fail in one of two general ways. On one hand, separability may fail due to a change in the salience of one dimension as a function of the level on the other dimension. Using our color/shape example, this would occur if circles and squares were harder to distinguish when they are purple than when they are red. In the static GRT model, this would produce, for example, higher d' for shape when the color is red than purple. On the other hand, separability may fail due to mean-shift integrality which we discuss in the next subsection. Mean shift integrality could lead, for example, to perceiving shapes as more square when they are red (whether they are actually squares or circles) and in which the shapes are perceived as more circular when they are purple. In this case, d' would be equal for shape, whether it is red or purple, though the underlying space would be warped by a shift toward square for red shapes and toward circle for purple shapes.

In this section we simulate salience-induced failure of separability in the dynamic GRT model, and in the next section we simulate mean-shift integrality.

Four salience-induced failure of separability models were simulated. In two, activation in the X channels was influenced by activation in the Y channels, but not vice-versa, depicted in Fig. 4. In the other two, depicted in Fig. 5, salience change was implemented symmetrically (i.e., salience on each dimension varied simultaneously as a function of the level on the other dimension). One of each of these models included input noise (i.e., non-zero $\boldsymbol{\eta}_a$), and one of each included perceptual noise (i.e., non-zero $\boldsymbol{\eta}_b$). The single-dimension salience-change model was implemented via the following matrix.

$$\mathbf{A}_{asym} = \begin{pmatrix} -0.005 & 0 & -0.010 & +0.010 \\ 0 & -0.005 & -0.010 & +0.010 \\ 0 & 0 & -0.005 & 0 \\ 0 & 0 & 0 & -0.005 \end{pmatrix}.$$

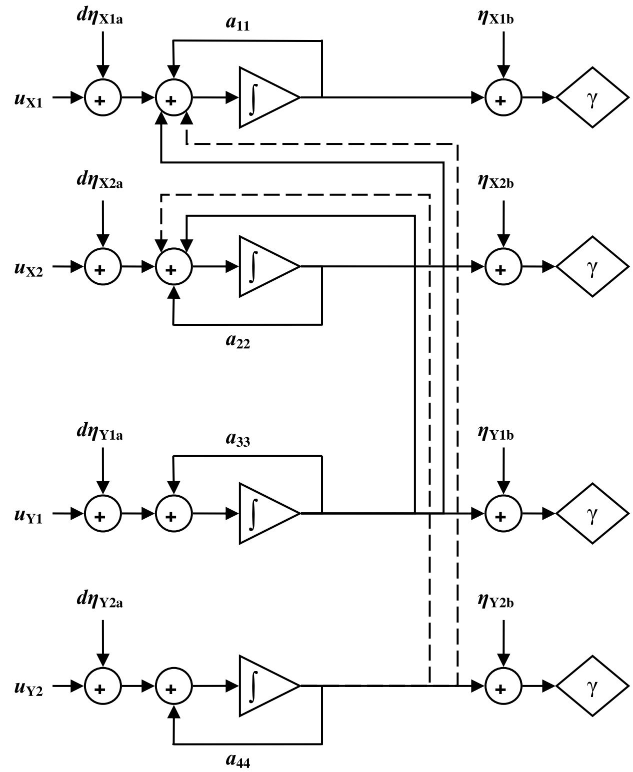

Fig. 4. Schematic representation of a four channel linear stochastic dynamic system with an asymmetric, salience-induced failure of separability. In this model, activation in the X channels was influenced by activation in the Y channel, but not vice-versa. Inhibitory connections, corresponding to negative values in the **A** matrix, are indicated by solid lines; facilitatory connections are indicated by dashed lines.

Here (as in all simulations reported here), the first row and column of the matrix correspond to X_1, the second to X_2, the third to Y_1, and the fourth to Y_2.

The non-zero off-diagonal values here caused the activation in both X channels to be suppressed by Y_1 and enhanced by Y_2. When input $s_y = 1$ was present, only noise was present in Y_2, which increased the activation in Y_1 and decreased the activation in each X channel. Increased activation due to input in either X channel was thereby counteracted, producing reduced accuracy and slower first passage times on the X channels. On the other hand, when input $s_y = 2$ was present, only noise was present in Y_1, and activation was increased in Y_2 and both X channels, producing a small increase in the activation of the X channel with no input and enhancing the input-driven activation in the X channel that received input. Hence, accuracy was increased and finishing times decreased, on average, in both X channels. The zeros in the bottom left of the matrix indicate that the level of the X dimension had no effect on the activation in either Y channel.

The symmetric salience change models used the following matrix:

$$\mathbf{A}_{sym} = \begin{pmatrix} -0.005 & 0 & -0.010 & +0.010 \\ 0 & -0.005 & -0.010 & +0.010 \\ -0.010 & +0.010 & -0.005 & 0 \\ -0.010 & +0.010 & 0 & -0.005 \end{pmatrix}.$$

In the symmetric model, as above, activation in the X channels was suppressed by activation in Y_1 and enhanced by activation in Y_2, and, simultaneously, activation in the Y channels was suppressed by activation in X_1 and enhanced by activation in X_2. In this case, we expect low accuracy and slow response times across the board in response to the $s_x = 1, s_y = 1$ stimulus, as activation is suppressed on both levels of each dimension. We

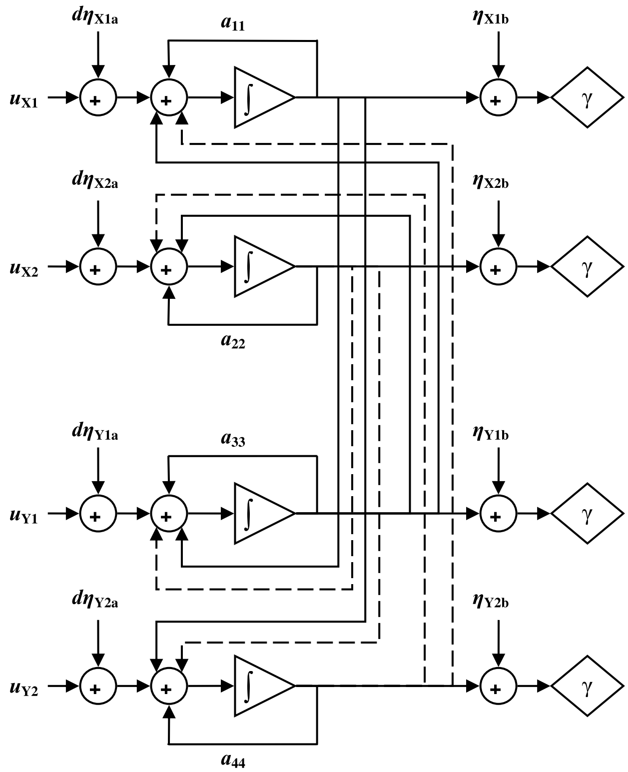

Fig. 5. Schematic representation of a four channel linear stochastic dynamic system with a symmetric, salience-induced failure of separability. In this model, activation in the X channels was influenced by activation in the Y channel and activation in the Y channels was influenced by activation in the X channel. Inhibitory connections, corresponding to negative values in the **A** matrix, are indicated by solid lines; facilitatory connections are indicated by dashed lines.

expect moderately low accuracy and moderately slow response times in response to the $s_x = 1, s_y = 2$ and $s_x = 2, s_y = 1$ stimuli, as activation is suppressed on one dimension and enhanced on the other, and we expect high accuracy and fast response times in response to the $s_x = 2, s_y = 2$ stimulus, as the inputs enhance activation on both dimensions.

Regardless of the type of noise, the speed of evidence accumulation in X_1 channel is higher when $s_y = 2$ so perceptual separability fails. When input noise is used, the cross-channel interactions will lead to stochastic dependence between the X and Y processes so perceptual independence fails. When there is only perceptual noise, there is no induced stochastic dependence, so perceptual independence holds.

By Proposition 3, we expect the violation of perceptual separability to induce failures of timed marginal response invariance. By Proposition 4, we also expect failures of static marginal response invariance. The asymmetric model should produce maximum differences (between defective CDFs) near zero for responses at both levels of s_y in essentially the same manner as the independent and separable model, but it should produce large negative differences for responses at both levels of s_x, since the difference for marginal R_x responses is derived from the defective CDF given $s_y = 1$ input, which produces lower accuracy and slower finishing times, minus the defective CDF given $s_y = 2$ input, which produces higher accuracy and faster finishing times. Following this logic, we expect the symmetric model to produce large negative differences, and large absolute differences, for both R_x and R_y responses.

When there is input noise, the cross-talk between channels will cause failure of perceptual independence, which should, in turn, induce failure of timed report independence by Proposition 5. It is somewhat more difficult to predict the pattern of the differences between empirical defective CDFs (and products thereof), and so the pattern of failures of timed report independence, in large part because it is difficult to estimate *a priori* the magnitude of the effect of multiplying marginal defective CDFs. We can make statements, however, regarding expected general differences between the asymmetric and symmetric models. Because, in the asymmetric models, Y_1 and Y_2 have blanket (negative and positive, respectively) effects on the X dimension, we expect similar timed report independence differences for the $s_x = 1, s_y = 1$ and $s_x = 2, s_y = 1$ stimuli (i.e., for the $s_y = 1$ stimuli), on the one hand, and for the $s_x = 1, s_y = 2$ and $s_x = 2, s_y = 2$ stimuli (i.e., the $s_y = 2$ stimuli), on the other. By way of contrast, in the symmetric models, we expect distinct differences for the $s_x = 1, s_y = 1$ stimulus, wherein both channels receiving input interact negatively with the other dimensions, for the $s_x = 2, s_y = 2$ stimulus, wherein both channels receiving input interact positively the other dimensions, and similar differences for the $s_x = 1, s_y = 2$ and $s_x = 2, s_y = 1$ stimuli, each of which consists of one negatively interacting and one positively interacting channel receiving input.

By Propositions 3 and 5, we expect rates of failure of *timed* marginal response invariance and *timed* report independence to be equal to or greater than rates of their static counterparts, with the reminder that, as in the baseline model, some violations of these implications can occur in the true absence of any underlying perceptual or decisional interactions.

The salience-induced failure of separability simulation results are shown in Table 3. The first, and most obvious, pattern in the timed marginal response invariance results (top four rows) is the expected failures of timed marginal response invariance and timed report independence, as well as the difference between the symmetric and asymmetric models; violations are essentially nonexistent on the Y dimension in the asymmetric model, whereas they occur (with varying rates) on the X dimension in this model and on both dimensions in the symmetric model. On the other hand, timed report independence (bottom four rows) is violated at comparable rates in both models, though not for all four stimuli.

Another notable difference is that between the input and perceptual noise simulations. Timed marginal response invariance is far more likely to be violated with perceptual noise (η_b) than with input noise (η_a). As a consequence of the dependence of perceptual independence on the type of noise, timed report independence should fail with input noise and not with perceptual noise. It is also noteworthy that timed marginal response invariance is more likely to be violated in the symmetric model when the stimuli are at level 2 than when they are at level 1 (i.e., when the inputs to one dimension enhance activation on the other rather than when they suppress it). Timed report independence is more likely to be violated in the $s_x = 1, s_y = 1$ and $s_x = 2, s_y = 2$ stimuli, and this effect is particularly pronounced in the symmetric model, which exhibits a sort of 'synergy' in these stimuli (i.e., negative or positive interactions, respectively, in both directions simultaneously) and a kind of counter-balancing between the positive and negative interactions in the $s_x = 1, s_y = 2$ and $s_x = 2, s_y = 1$ stimuli.

The static marginal response invariance results do not track the timed results very closely. Although violations of timed marginal response invariance are frequent in the symmetric model with perceptual noise, violations of static marginal response invariance occur very rarely for all combinations of model symmetry and noise type. Note, though, that the rates of failure of timed marginal response invariance exceed the rates of failure of static marginal response invariance in all but a few cases, which is consistent with, if not directly supportive of, Proposition 2.

The timed and static report independence results correspond more closely to one another, with the timed results showing slightly higher rates of violation than the static, consistent with Proposition 6. Hence, we might reasonably expect the timed tests to be as or more sensitive to underlying salience-induced interactions than are the static tests.

Table 3
Percentage of simulations with significant results for salience change models with input noise (η_a) or perceptual noise (η_b). In these models perceptual separability fails, so static marginal response invariance should fail. When the symmetry condition holds, timed marginal response invariance should fail; the theory above does not dictate the outcome when the symmetry condition fails. Perceptual independence holds when there is only perceptual noise, but fails when there is input noise, and timed and static report independence should follow that pattern. The predicted failures are indicated in bold. In each case the timed tests are either more sensitive or about equally sensitive to salience changes.

	Constant	Timed				Static			
		Asymmetric		Symmetric		Asymmetric		Symmetric	
		η_a	η_b	η_a	η_b	η_a	η_b	η_a	η_b
MRI	$R_x = 1; s_x = 1$	0.09	0.82	**0.09**	**0.35**	**0.06**	**0**	**0.06**	**0**
	$R_x = 2; s_x = 2$	0.08	0.80	**0.12**	**0.99**	**0.08**	**0**	**0.05**	**0**
	$R_y = 1; s_y = 1$	0	0	**0.05**	**0.44**	**0.03**	**0.01**	**0.08**	**0**
	$R_y = 2; s_y = 2$	0.01	0	**0.16**	**0.98**	**0.04**	**0.04**	**0.08**	**0**
RI	$s_x = 1; s_y = 1$	**0.21**	0.02	**0.48**	0.01	**0.17**	0.07	**0.34**	0.05
	$s_x = 1; s_y = 2$	**0.15**	0.02	**0.05**	0.01	**0.14**	0.04	**0.05**	0.04
	$s_x = 2; s_y = 1$	**0.11**	0.03	**0.06**	0.01	**0.11**	0.06	**0.05**	0.04
	$s_x = 2; s_y = 2$	**0.21**	0.05	**0.27**	0.06	**0.18**	0.06	**0.22**	0.07

3.4. Failure of separability due to pseudo-mean-shift integrality

Mean-shift integrality is a way that perceptual separability fails even if there is no change in the salience of one dimension as a function of the other dimension. By way of illustration, reconsider our shape-color example, in which the stimuli and responses consist of red and purple squares and rectangles. In this context, the presence of mean shift integrality could mean that the shape of the stimuli is perceived as more rectangular when purple than when red, that the color of the stimuli is perceived as more purple when rectangular than when square, or both. Moreover, in the static GRT, both of the marginal distributions for color are shifted by the same amount when the experimenter presents a rectangle instead of a square, and/or both of the marginal distributions for shape are shifted by the same amount when the experimenter presents a purple stimulus instead of a red stimulus. This shift results in an invariance of d' on color over the change in shape. Of course, it is possible to have this type of invariance on one dimension and not the other.

Even in the static case, although the d's for the separate settings of the alternate dimension are equal, scrutiny of the hit and false alarm frequencies can readily detect such shifts, as long as decisional separability is not violated in exactly the manner needed to keep them constant.

When we move to the present dynamic theory, there is very different and somewhat more complex mathematical structure. Notice that the notion of a single, unique marginal distribution does not even exist. For a first start on this issue, we chose to implement a type of pseudo-mean-shift integrality which is elicited through dynamic interactions of the parallel channels. Analogous to the salience interactions above, increases or decreases on one dimension can feed into the input or, alternatively, the outgoing perceptual conduit of the other dimension. Consider for instance, the case where the Y channel affects the X channel. Suppose Y_1 facilitates X_1 but inhibits X_2 and Y_2 facilitates X_2 but inhibits X_1. Since the parameters of interaction will be the same for either X_1 or X_2, the help or hurt to the X channel is invariant for X_1 and X_2, thereby producing a type of shift-integrality. However, we will employ the term "pseudo-mean-shift integrality" as a reminder of the theoretical context.

Four pseudo-mean-shift integrality models were simulated. As in the salience-change models, two of the pseudo-mean-shift integrality models were asymmetric such that activation in the X channels was influenced by activation in the Y channels, but not vice-versa. This structure is depicted in Fig. 6. In the other two, depicted in Fig. 7 mean-shift integrality was implemented

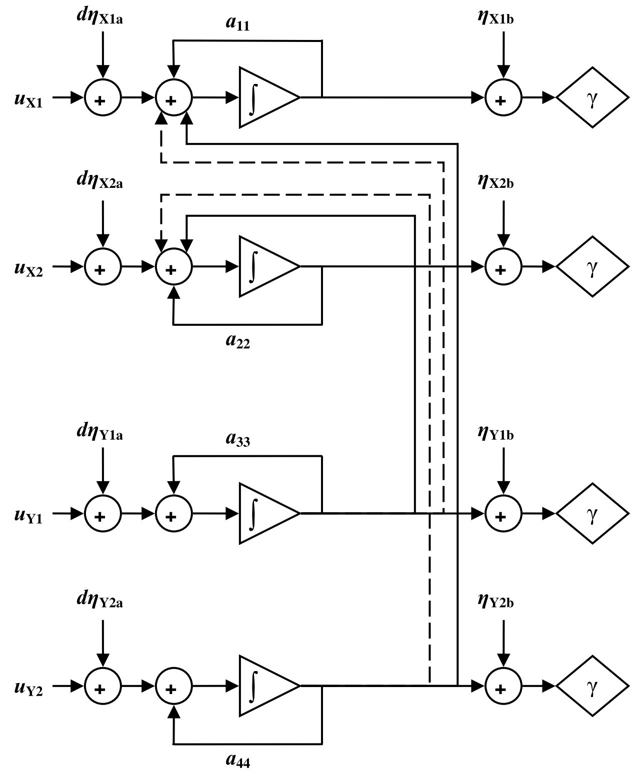

Fig. 6. Schematic representation of a four channel linear stochastic dynamic system with an asymmetric interactions and pseudo-mean-shift integrality. In this model, activation in the X channels was influenced by activation in the Y channel, but not vice-versa. Inhibitory connections, corresponding to negative values in the **A** matrix, are indicated by solid lines; facilitatory connections are indicated by dashed lines. Notice that this model differs from the salience-induced change models by which connections are facilitatory and inhibitory.

symmetrically such that the X channels influenced the Y channels and vice versa in essentially the same way. Also as above, one of each of these models included input noise, and one of each included perceptual noise. These models have the same pattern of violations of perceptual separability and perceptual independence as the salience-change models. Perceptual separability is always violated; perceptual independence is violated with input noise but not perceptual noise. Hence, we expect timed marginal response invariance should fail (Proposition 3) and time report independence should only hold when there is no input noise (Proposition 5).

The asymmetric mean-shift model was implemented via the following matrix:

$$\mathbf{A}_{asym} = \begin{pmatrix} -0.005 & 0 & +0.010 & -0.010 \\ 0 & -0.005 & -0.010 & +0.010 \\ 0 & 0 & -0.005 & 0 \\ 0 & 0 & 0 & -0.005 \end{pmatrix}.$$

Whereas, as noted earlier, in static GRT, mean-shift integrality is defined in terms of the relative locations of the means of perceptual distributions, we defined pseudo-mean-shift integrality as a function of the mean activation in the processing channels. Here, activation in X_1 is enhanced by activation in Y_1 and suppressed by activation in Y_2, whereas activation in X_2 is suppressed by activation in Y_1 and enhanced by activation in Y_2. Thus, when $s_y = 1$ input was present, $R_x = 1$ responses were more likely and occurred more quickly, whether correct or incorrect, whereas when $s_y = 2$ input was present, $R_x = 2$ responses were more likely and more rapid, whether correct or incorrect. The symmetric pseudo-mean-shift model added an analogous set of interactions

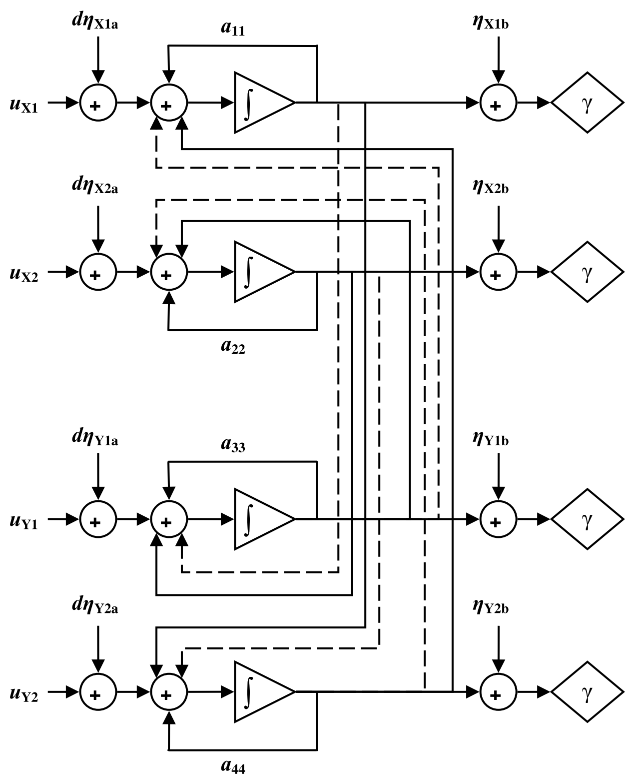

Fig. 7. Schematic representation of a four channel linear stochastic dynamic system with an asymmetric interactions and pseudo-mean-shift integrality. In this model, activation in the X channels was influenced by activation in the Y channel and activation in the Y channels was influenced by activation in the X channel. Inhibitory connections, corresponding to negative values in the **A** matrix, are indicated by solid lines; facilitatory connections are indicated by dashed lines. Notice that this model differs form the salience-induced change models by which connections are facilitatory and inhibitory.

causing activation in the Y channels to be influenced by activation in the X channels.

The symmetric pseudo-mean-shift model was implemented via the following matrix:

$$\mathbf{A}_{symm} = \begin{pmatrix} -0.005 & 0 & +0.010 & -0.010 \\ 0 & -0.005 & -0.010 & +0.010 \\ +0.010 & -0.010 & -0.005 & 0 \\ -0.010 & +0.010 & 0 & -0.005 \end{pmatrix}.$$

As with the salience-induced failures of separability (and independence), we expect these channel interactions to induce failures of timed and static marginal response invariance (Propositions 3 and 4) and timed report independence (Proposition 5). We also expect the implications of Propositions 2 and 6 to allow for, as above, higher rates of failure of the timed tests than their static counterparts.

Simulation results are shown in Table 4. As with salience-induced failure of separability, we see numerous failures of timed marginal response invariance and timed report independence, as well as the expected difference between the symmetric and asymmetric models with respect to violations of timed marginal response invariance. Violations are at or below 0.05 for responses on the Y dimension for the asymmetric model and well above 0.05 for responses on X for the asymmetric model and for responses on both dimensions for the symmetric model. Violations of timed report independence are very frequent with input noise and very infrequent with perceptual noise for both the asymmetric and symmetric models.

Overall rates of violation of both statistics is much higher for the pseudo-mean-shift failures of separability than it is for the salience-induced failures, though it is not immediately evident why (e.g., the magnitude of the interactions between channels is identical in the two sets of models, and all other parameters are held constant). Violations occurred in all (or nearly all) simulations for some combinations of symmetry and noise type.

Table 4
Percentage of simulations with significant results for pseudo-mean-shift integrality models with input noise (η_a) or perceptual noise (η_b). In these models perceptual separability fails, so static marginal response invariance should fail. When the symmetry condition holds, timed marginal response invariance should fail; the theory above does not dictate the outcome when the symmetry condition fails. Perceptual independence holds when there is only perceptual noise, but fails when there is input noise, and timed and static report independence should follow that pattern. The predicted failures are indicated in bold.

	Constant	Timed				Static			
		Asymmetric		Symmetric		Asymmetric		Symmetric	
		η_a	η_b	η_a	η_b	η_a	η_b	η_a	η_b
MRI	$R_x = 1; s_x = 1$	0.99	1.00	**0.45**	**1.00**	**1.00**	**1.00**	**0.81**	**1.00**
	$R_x = 2; s_x = 2$	0.98	1.00	**0.35**	**1.00**	**0.00**	**0.00**	**0.00**	**0.00**
	$R_y = 1; s_y = 1$	0.05	0	**0.56**	**1.00**	**0.08**	**0.01**	**0.84**	**1.00**
	$R_y = 2; s_y = 2$	0.04	0	**0.41**	**1.00**	**0.02**	**0.06**	**0.00**	**0.00**
RI	$s_x = 1; s_y = 1$	**1.00**	0.01	**1.00**	0	**1.00**	0.07	**1.00**	0.02
	$s_x = 1; s_y = 2$	**1.00**	0.02	**1.00**	0	**1.00**	0.04	**1.00**	0.04
	$s_x = 2; s_y = 1$	**1.00**	0.04	**1.00**	0.05	**1.00**	0.04	**1.00**	0.07
	$s_x = 2; s_y = 2$	**1.00**	0	**1.00**	0	**1.00**	0.04	**1.00**	0.04

As in the salience-induced failure of separability simulations, the pseudo-mean-shift simulations produced some cases in which static marginal response invariance is not violated while the timed analog shows extensive violations (e.g., for $R_y = 2$, $s_y = 2$ in the symmetric model and for $R_x = 2$; $s_x = 2$ in both models). Unlike the previous simulations, however, there are also pseudo-mean-shift cases in which static marginal response invariance is violated more often than its timed analog (e.g., with $R_x = 1$, $s_x = 1$ and $R_y = 1$, $s_y = 1$ with input noise), contra Proposition 2.

The timed and static report independence tests correspond more closely to one another. Both show violations in 100% of simulations with input noise, and both show very few violations with perceptual noise which follows the predictions based on the effect of noise type on perceptual independence. Allowing for the kind of minor violations of Proposition 6 that are evident even in the baseline model, these report independence results are consistent with the expected implicational relationship between timed and static statistics.

3.5. Failure of perceptual independence and presence of perceptual separability

Simulations were run with four models exhibiting failure of perceptual independence and presence of perceptual separability. In these models, the channels did not interact, as interaction has been described thus far (i.e., the **A** matrix was proportional to the identity matrix; depicted in Fig. 1). Rather, the input and perceptual noise were based on multivariate normal random variables with the following correlation matrices. In two of the "pre-correlated" noise models, input noise was added ($\boldsymbol{\eta}_a$), and in the other two perceptual noise was added ($\boldsymbol{\eta}_b$).

Because perceptual separability holds in these models, we expect timed marginal response invariance to hold (Proposition 3). Because perceptual and decisional separability hold in this model, and because we expect timed marginal response invariance to hold, we also expect static marginal response invariance to hold (Propositions 2 and 4). On the other hand, we expect the correlations in the noise to induce failure of perceptual independence, which should in turn induce failures of timed report independence (Proposition 5). Rates of failure of timed report

Table 5
Percentage of simulations with significant results independent models with positively or negatively correlated input with either input noise (η_a) or perceptual noise (η_b). In these models, perceptual independence fails so timed and static report independence should fail. Perceptual separability holds so static marginal response invariance should hold and timed marginal response invariance should hold as long as the symmetry condition holds; the theory above does not dictate the outcome when the symmetry condition fails. The predicted failures are indicated in bold.

	Constant	Timed				Static			
		+Correlated		−Correlated		+Correlated		−Correlated	
		η_a	η_b	η_a	η_b	η_a	η_b	η_a	η_b
MRI	$R_x = 1; s_x = 1$	0	0	0.01	0	0.02	0.00	0.03	0.01
	$R_x = 2; s_x = 2$	0	0	0.01	0	0.01	0.01	0.01	0.07
	$R_y = 1; s_y = 1$	0	0.02	0.01	0	0.02	0.02	0.03	0.02
	$R_y = 2; s_y = 2$	0	0	0	0	0.05	0.00	0.03	0.02
RI	$s_x = 1; s_y = 1$	**0.08**	**0.05**	**0.11**	**0.08**	**0.06**	**0.03**	**0.04**	**0.05**
	$s_x = 1; s_y = 2$	**0.07**	**0.01**	**0.07**	**0.02**	**0.07**	**0.07**	**0.07**	**0.03**
	$s_x = 2; s_y = 1$	**0.03**	**0**	**0.06**	**0.04**	**0.04**	**0.04**	**0.07**	**0.08**
	$s_x = 2; s_y = 2$	**0.02**	**0.03**	**0.07**	**0.07**	**0.01**	**0.07**	**0.04**	**0.06**

independence should not be (substantially) less than rates of failure of static report independence (Proposition 6).

$$\mathbf{R}_{neg.} = \begin{pmatrix} 1.000 & -0.325 & -0.325 & -0.325 \\ -0.325 & 1.000 & -0.325 & -0.325 \\ -0.325 & -0.325 & 1.000 & -0.325 \\ -0.325 & -0.325 & -0.325 & 1.000 \end{pmatrix}$$

$$\mathbf{R}_{pos.} = \begin{pmatrix} 1.000 & 0.325 & 0.325 & 0.325 \\ 0.325 & 1.000 & 0.325 & 0.325 \\ 0.325 & 0.325 & 1.000 & 0.325 \\ 0.325 & 0.325 & 0.325 & 1.000 \end{pmatrix}.$$

Simulation results are listed in Table 5. Neither timed marginal response invariance nor timed report independence were violated with any regularity in any of these models. Of course, the lack of violations of (timed or static) marginal response invariance was expected, as the correlated noise simulations were run with models known to be non-interactive so that perceptual separability held.

The lack of violations of (timed or static) report independence in these models is somewhat surprising. The magnitude of the channel interactions is the same as in the salience-induced and mean-shift failure of separability models discussed above, and each of these produced numerous violations of report independence (and marginal response invariance), and the multivariate noise was correlated at fairly high levels. One possibility is that the statistical tests are simply not powerful enough to detect the violations. No matter what the source of this failure, it is an important topic of future exploration.

4. Empirical application

We finish with a brief presentation of the application of the results presented above to data collected in two speech perception experiments. Identification data from two experiments probing the perception of the English consonants '*p*', '*b*', '*t*', and '*d*' in different syllable positions were analyzed by testing static marginal response invariance and report independence (Silbert, 2010). The stimuli in these experiments were defined on the linguistic dimensions 'place of articulation' and 'voicing', with '*p*' and '*b*' being X_1 = 'labial', '*t*' and '*d*' being X_2 = 'alveolar', '*p*' and '*t*' being Y_1 = 'voiceless', and '*b*' and '*d*' being Y_2 = voiced. In experiment one, these four consonants occurred in onset position in the nonsense syllables '*pa*', '*ba*', '*ta*', and '*da*', and in experiment two they occurred in coda position in the nonsense syllables '*ap*', '*ab*', '*at*', and '*ad*' (all vowels were '*a*' as in 'father'). Silbert (2010) reports that, for consonants in onset position, marginal response invariance fails across the board for $R_x = 2$, tends to fail for $R_y = 2$, tends to hold for $R_y = 1$ and $R_x = 1$, and perceptual independence tends to hold with few exceptions (e.g., $s_x = 2$, $s_y = 2$, or '*d*'). On the other hand, marginal response invariance and report independence both tend to hold for these consonants in coda position. In both experiments, the static analyses provide support for decisional separability.

To demonstrate the use of the new methods established in this paper, we apply them to the data from Silbert (2010). The results of the first experiment are shown in Tables 6 and 7. The results of the second experiment are shown in Tables 8 and 9.

Unlike with the simulations described above, with the empirical application, we do not know *a priori* which, if any, of the assumptions underlying our theoretical developments hold. For example, it may be that perceptual separability, perceptual independence, and/or decisional separability fail for (some of) the participants in these experiments. And although we consider it extremely unlikely, it may be that identification of the phonological features under consideration here does not occur in parallel. If we allow ourselves the assumption of parallel processing with decisional separability, we can still probe the experimental data for the implicational relationships between the timed and static statistics of interest.

In the onset data, there are 21 failures of timed marginal response invariance and 19 failures of static marginal response invariance. By and large, failures of the former correspond to failures of the latter, though there are five cases in which a timed test fails while a static test does not (e.g., Participant 1, $R_x = 1, s_x = 1$) or vice versa (e.g., Participant 2, $R_x = 2, s_x = 2$), which is consistent with Proposition 2, and three cases in which the static test fails while the timed does not, inconsistent with Proposition 2. For both timed and static tests, most failures occur for $R_x = 2, s_x = 2$ and $R_y = 2, s_y = 2$, consistent with the results reported by Silbert (2010), indicating that '*d*' is more perceptually distinct from '*b*' than '*t*' is from '*p*' and that '*d*' is more perceptually distinct from '*t*' than '*b*' is from '*p*'.

There are no failures of timed report independence in this data set, despite the fact that there are 9 failures of static report independence, inconsistent with Proposition 6. There is no obvious pattern in the report independence results. Half of the participants exhibit failures for '*d*', and the other 5 failures occur for various combinations of participant and stimulus.

In the coda data, there are 16 failures of timed marginal response invariance and 12 failures of static marginal response invariance. As in the onset data, the two sets of results track one another fairly closely, though there are cases of divergence (e.g., Participant 1, $R_y = 1, s_y = 1$ and $R_y = 2, s_y = 2$), seven of which are consistent with Proposition 2, three of which are not. As in the results reported by Silbert (2010), the pattern indicating high perceptual distinctiveness for '*d*' in syllable onsets does not occur in syllable codas.

There is no clear and consistent pattern of failure of marginal response invariance in this data set. There are 2 failures of timed report independence and 8 failures of static report independence in the data from the second experiment. As with the onset data, this represents a violation of Proposition 6. Both of the timed report independence failures correspond to static report independence failures, though, again, there is no clear and consistent pattern to the failures of report independence.

The overall pattern of results suggests that tests of timed and static marginal response invariance are roughly equally statistically sensitive, though it seems that tests of timed report independence are less sensitive than are the corresponding static tests. Based on Proposition 6, whenever static report independence fails, timed report independence should also fail. Thus the difference in sensitivity between the two measures is due to the statistical test being used. We will return to the issue of statistical testing in the discussion.

Table 6
Tests of timed and static marginal response invariance for response time data from Experiment 1 of Silbert (2010), probing identification of voicing and place of articulation in syllable onset position. With $\alpha = 0.05$, the critical value of the KS statistic for timed marginal response invariance tests is 1.36. The critical value of the z statistics for static marginal response invariance is 1.96.

		Participant							
		1	2	3	4	5	6	7	8
Timed MRI	$R_x = 1; s_x = 1$	**1.932**	**2.09**	1.307	0.704	1.255	**2.839**	**2.075**	0.902
	$R_x = 2; s_x = 2$	**4.198**	**3.88**	**4.409**	**2.559**	**3.855**	**3.489**	**5.253**	**4.333**
	$R_y = 1; s_y = 1$	0.541	0.934	0.712	**1.47**	**1.481**	0.987	**1.633**	0.774
	$R_y = 2; s_y = 2$	**1.925**	**2.481**	**1.93**	**2.889**	**4.787**	1.079	**2.283**	0.963
Static MRI	$R_x = 1; s_x = 1$	1.519	−**2.612**	−0.699	0.716	**2.008**	−**2.772**	−**3.035**	1.536
	$R_x = 2; s_x = 2$	−**11.197**	−**7.138**	−**6.86**	−**7.28**	−**8.365**	−**6.196**	−**7.957**	−**12.857**
	$R_y = 1; s_y = 1$	−0.272	−**3.537**	−1.833	−0.991	−1.834	−1.223	−0.674	0.797
	$R_y = 2; s_y = 2$	−**2.446**	−**2.288**	−**2.155**	−**7.676**	−**3.853**	−**2.296**	−1.525	−1.708

Table 7
Tests of timed and static report independence for response time data from Experiment 1 of Silbert (2010), probing identification of voicing and place of articulation in syllable onset position. With $\alpha = 0.05$, the critical value for the test of timed report independence is 0.3. The critical value for the test of static report independence is 3.84.

		Participant							
		1	2	3	4	5	6	7	8
Timed RI	$s_x = 1; s_y = 1$	0.023	0.109	0.061	0.046	0.059	0.057	0.037	0.076
	$s_x = 1; s_y = 2$	0.029	0.168	0.066	0.155	0.217	0.063	0.015	0.07
	$s_x = 2; s_y = 1$	0.122	0	0.031	0.057	0.014	0.069	0.095	0.062
	$s_x = 2; s_y = 2$	0	0	0.012	0.16	0.003	0.096	0	0.266
Static RI	$s_x = 1; s_y = 1$	0.111	2.59	0.388	0.795	1.088	1.461	0.53	**4.182**
	$s_x = 1; s_y = 2$	0.205	**25.37**	0.227	2.472	**6.748**	3.172	0.027	1.38
	$s_x = 2; s_y = 1$	**5.724**	0.126	0.011	0.758	0.266	3.528	**6.668**	1.927
	$s_x = 2; s_y = 2$	2.927	**5.224**	0.002	**39.489**	0.391	**62.655**	2.927	**68.7**

Table 8
Tests of timed and static marginal response invariance for Experiment 2 of Silbert (2010), probing identification of voicing and place of articulation in syllable coda position. With $\alpha = 0.05$, the critical value of the KS statistic for timed marginal response invariance tests is 1.36. The critical value of the z statistics for static marginal response invariance is 1.96.

		Participant							
		1	2	3	4	5	6	7	8
Timed MRI	$R_x = 1; s_x = 1$	**3.04**	0.6	**2.424**	**1.857**	**1.369**	**1.581**	1.047	0.747
	$R_x = 2; s_x = 2$	**1.715**	**1.49**	1.297	1.056	0.782	**3.574**	**2.595**	**1.829**
	$R_y = 1; s_y = 1$	**2.045**	1.194	0.598	**1.758**	0.877	**1.997**	**1.586**	**2.481**
	$R_y = 2; s_y = 2$	0.687	0.953	0.797	1.248	0.815	1.298	**1.859**	0.949
Static MRI	$R_x = 1; s_x = 1$	**6.524**	1.56	**3.141**	**4.905**	1.013	−**3.262**	−0.102	−0.831
	$R_x = 2; s_x = 2$	−**4.06**	1.744	−1.635	−**2.535**	−0.401	**2.917**	**2.513**	−1.339
	$R_y = 1; s_y = 1$	−1.159	0.175	0.503	−**2.82**	−0.897	−0.709	0.357	−**2.288**
	$R_y = 2; s_y = 2$	−**3.158**	−1.831	1.508	−**2.027**	−0.727	−0.709	−1.688	−0.709

Table 9
Tests of timed and static report independence for Experiment 2 of Silbert (2010), probing identification of voicing and place of articulation in syllable coda position. With $\alpha = 0.05$, the critical value for the test of timed report independence is 0.3. The critical value for the test of static report independence is 3.84.

		Participant							
		1	2	3	4	5	6	7	8
Timed RI	$s_x = 1; s_y = 1$	0.017	0.112	0.009	0.138	0.041	0.016	0.017	0.065
	$s_x = 1; s_y = 2$	0.223	**0.346**	0	0.128	0.055	0.01	0.069	0.01
	$s_x = 2; s_y = 1$	0	**0.343**	0.059	0.008	0.021	0	0.032	0
	$s_x = 2; s_y = 2$	0.01	0.079	0.056	0.181	0.044	0	0.076	0
Static RI	$s_x = 1; s_y = 1$	0.125	1.993	0.011	**6.429**	**4.623**	0.281	0.028	1.563
	$s_x = 1; s_y = 2$	**9.376**	**8.891**	0.057	**8.49**	1.314	0.031	1.17	0.018
	$s_x = 2; s_y = 1$	0.015	**16.729**	3.613	0.022	2.807	0.264	0.436	0.089
	$s_x = 2; s_y = 2$	0.031	0.128	2.689	**14.394**	**14.554**	0.061	2.803	0.174

5. Discussion

5.1. Analytic theorems, simulations with noisy dynamic systems and initial data

This project extended General Recognition Theory to a time-dynamic, parallel-channels, theory. The new theoretical structures entailed expanded definitions of the various types of independence as well as novel combined time-accuracy versions of key testable statistics such as marginal response invariance and sampling (report) independence. We employed this theory to interrelate response times and deduce theorems on the various types of independence which combine these observable variables. Interestingly, analogs to the static theory exist in each case, although a new statistic, timed marginal response invariance required extra conditions to hold for the predicted equality to hold. Also in general, when the timed versions of the predicted (on the basis of the various types of independence) relationships hold, the static predictions will also hold, although the reverse is not true.

Within the generalized theory, we discovered that certain concepts are more closely connected than in the static version. An especially salient example is the relationship between perceptual independence and perceptual separability. These concepts are quite distinct in the original theory. However, in the new theory, failure of perceptual independence through interconnections of the two major channels, will typically cause failures in perceptual separability

In order to augment the analytic relationships between theory and data, we simulated data for a variety of stochastic linear dynamic models. This enabled an exploration of the empirical implications of (a) two ways in which noise may enter an information processing system, and (b) various ways in which processing channels may interact. The results of the simulations suggest that distinct observable statistics provide substantial fine-grained information about the properties of the underlying system. For example, the nature of underlying failure of perceptual separability (e.g., salience-induced failure vs. mean-shift integrality, asymmetric interactions vs. symmetric interactions) produce distinct patterns of presence or failure of timed marginal response invariance and timed report independence across the four stimuli and responses defined in the basic factorial GRT experimental protocol.

Finally, a brief illustration of our stochastic GRT approach was given as applied to data collected in two speech perception experiments (reported in Silbert, 2010). This illustration provides both reason for confidence in the approach and points the way toward a number of future directions for related work. The new results for the 'place by voicing' experiment of Silbert (2010) are largely consistent with the results reported there—with numerous failures of perceptual separability and a smaller number of failures of independence.

5.2. Relationship to other theories

This work extends Ashby's (1989, 2000) dynamic systems approach to GRT. In the first of these papers, a stochastic GRT was constructed through a class of discrete time linear systems models, whereas the second paper uses continuous stochastic linear systems. While Ashby's stochastic GRT model is shown to encompass the static model (see, e.g., Ashby (1989, pp. 447–449)), the dynamic model produces response time predictions only for a two-choice categorization experimental paradigm, and then only by virtue of the fact that the model implements a multidimensional random walk with absorbing barriers. Ashby also suggested a first-passage decision rule in circumstances of speed stress. In the absence of speed stress, he suggested that the observer waits until a point of maximum activation at which time she/he locates the closest boundary and makes the appropriate decision (also see Ashby & Maddox, 1991, 1994; Maddox & Ashby, 1996). By way of contrast, our model produces response time predictions for a full-factorial GRT paradigm via multiple, parallel evidence accumulators. Because our model can be restricted to form a two-choice, multidimensional random walk, it is a generalization of Ashby's model.

Ashby's second paper constructed a continuous-time stochastic GRT model for two stimuli (plus noise) and responses. As in our present developments, he assumed the existence of decision bounds that when reached, would precipitate a decision and response. This led immediately to a diffusion process. An important theorem was proven which relates predictions in the stochastic case to those emanating from the static GRT model in addition to other results. This model was not applied to assessment of perceptual independence. Our investigated set of stochastic parallel systems is entirely general, in the sense that it subsumes virtually any kind of parallel models which assume that the various channels race to individual criteria, in order to determine the decision. Even the broad class of linear or even non-linear stochastic systems is subsumed in our general class. However, stochastic linear systems which possess rich potential for exploration of channel interactions, effects of signal and noise characteristics and decisional variables were used to illustrate certain features of our theory.

Due in large part to the close theoretical links between the present work and the work reported in Ashby (1989, 2000), there are also more or less direct theoretical (and historical) links between the present work and that described by Ashby and Maddox (1994). However, there are also important differences. Ashby and Maddox (1994) employ the RT-Distance hypothesis in a study of the relationships between General Recognition Theory and so-called 'Garner' integrality—interference and redundancy effects in a particular kind of classification task. One crucial difference is our focus on perceptual relationships between dimensions in a factorial identification paradigm, whereas Ashby and Maddox (1994) are concerned with two-choice identification and classification tasks. It is also important to note that the RT-Distance hypothesis is essentially descriptive, as it "makes no processing assumptions" (Ashby & Maddox, 1994, p. 438), whereas our model is explicitly aimed at characterizing the information accumulation process and its relationship to various response-time-based statistics. Furthermore, the RT-Distance hypothesis assumes (or is agnostic with respect to the idea) that a static (variable) perceptual space drives categorization times; in the present work, we assume that response times *and* static, variable perceptual space are produced by a parallel-channel information processing system. Finally, we suspect that the parallel system described herein predicts RT-Distance effects, at least in certain cases, though it is beyond the scope of this paper to establish the mathematical relationships between our model and the RT-Distance hypothesis in any detail. We will return to the relationship of our results to Ashby's earlier findings in the Discussion.

The present work also bears at least superficial similarity to some of the work by Lamberts and colleagues (e.g., Lamberts, 2000; Lamberts et al., 2003). Again, though, there are crucial differences. For example, Lamberts' Extended Generalized Context Model (EGCM) is a model of the cognitive mechanisms driving two-choice categorization response times. On the other hand, our model is intended to characterize the process of information accumulation in, as noted repeatedly above, a factorial identification paradigm. In addition, in the EGCM, as described in Lamberts (2000), evidence in favor of one response alternative is, by necessity, evidence against the other alternative. In the parallel system described here, information accumulation in one channel need not imply the opposite in another channel, though particular kinds of interactions between channels may produce such behavior in special cases. We suspect that our process model is consistent (or at least not totally inconsistent) with the EGCM's hypothesized mechanisms, though, again, it is beyond the scope of this paper to establish this one way or the other. It is interesting to note that, if, in fact, our model implies the RT-Distance hypothesis, then to the extent that the RT-Distance hypothesis is inconsistent with the EGCM, there is a conflict between the EGCM model and the model described here.

In order to augment the analytic relationships between theory and data, we simulated data for a variety of stochastic linear dynamic models. This enabled an exploration of the empirical implications of (a) two ways in which noise may enter an information processing system, and (b) various ways in which processing channels may interact. The results of the simulations suggest that distinct observable statistics provide substantial fine-grained information about the properties of the underlying system. For example, the nature of underlying failure of perceptual separability

(e.g., salience-induced failure vs. mean-shift integrality, asymmetric interactions vs. symmetric interactions) produce distinct patterns of presence or failure of timed marginal response invariance and timed report independence across the four stimuli and responses defined in the basic factorial GRT experimental protocol.

5.3. Benefits, limitations and cautions, and future directions

A considerable portion of our past labor has focused on the development of theory-driven methodologies that allow the assessment of critical mechanisms and their functioning, that go beyond specific parameterized models. To date, most of these have been limited to RT or accuracy alone. Methods and models of RT that do not predict inaccuracies are open to criticism, due to potential disregard of such matters as speed–accuracy trade-offs. However, in our opinion, it is equally just to raise analogous questions when models or methods of accuracy ignore RTs, although for some reason this apprehension is not seen as frequently in the literature.

The present approach offers at least the following benefits: 1. The general framework provides a broad setting for the development and testing of dynamic models. Such models are thereby called upon to delineate the hypothesized mechanisms at a finer-grain level than that provided by an accuracy-based, or response time-based account alone. 2. In particular by doing so, response times and accuracy are amalgamated to further constrain and test models of classification. 3. Speed–accuracy relationships become open to inspection. 4. The non-parametric, distribution free approach, like static General Recognition Theory, tests vast classes of models rather than a single parameterized model.

Consider a single example which captures some of the last paragraph: Two data conditions with accuracy alone might suggest that because of equal hits and false alarms, perceptual signal-to-noise ratio is identical in the two conditions. However, if it were determined that the second condition evoked much longer response times, this inference would likely be put aside. Obviously, such findings can be used to test alternative dynamic models. Comparable aspects can be found in the simulation results as well as the subsequent experimental data.

From one point of view, the use of parallel architectures can be interpreted as a restriction, since our theorems do not include all architectures. Yet, when the three key features of perceptual independence, perceptual separability, and decisional separability are interpreted within, say serial systems, we do not expect the time-accuracy data predictions to remain unchanged. In fact, we expect the distinct predictions of independence/dependence statistics to assist in determining both the architecture as well as the failures of the key features. Thus, it is natural to consider architectures one at a time in this regard.

A definite limitation is the present restriction to decisionally separable models. In the original appearance of General Recognition Theory (Ashby & Townsend, 1986), the major theorems established sufficient conditions (on the types of independence) to ensure observable outcomes in the form of marginal response invariance and sampling (report) independence, and subsequently, the Gaussian-based theory (Kadlec & Townsend, 1992b). It was easy to find counterexamples which demonstrated the importance of decisional separability in such predictions. We strongly believe that an analogous situation holds here, and that in the absence of (the new definition of) decisional separability, such regularities as timed report independence fail. However, the companion counterexamples have not yet been developed.

It seems likely that the most publicized apprehension concerning failure of decisional separability is mean shift integrality. This is a theory-bound concept since it depends on model and perhaps distributional assumptions. Although these are rarely spelled out in detail, it appears that necessary and sufficient conditions for two-dimensional joint distributions should be that the marginal distributions on one (or the other or both) dimensions are identical, except for a shift in mean. Because of the interplay of the dynamic analogues to d' (e.g., drift rates), with the decisional criteria, it is far from transparent what if any analogues of mean shift integrality exist. We are presently commencing effort on this particular issue, and failure of decisional separability in general. Naturally, there remains much work to be done in this direction, as one of the great strengths of GRT is the conceptual distinction between perceptual and decisional interactions. Until more is known about this aspect of the theory, caution is advised on the part of researchers.

Another issue that warrants further investigation is that of statistical testing. To apply the new theory to data, some variety of statistical measure is necessary. We have included two simple statistical tests to illustrate the use of the theory, but hypothesis testing is not the focus of this paper. In future work, it will be important to more thoroughly investigate the properties of these statistical tests as alternatives. Better statistical tests are clearly needed, particularly for timed reported independence, given the results of the Empirical Application section.

Although the initial experimental applications are quite favorable, it would be well to affirm the workability of the new theory with experiments on "tried-and-true" sets of dimensions which have, so far at least, been found to be perceptually independent (separable, etc.) vs. sets which have not. For instance, Fifić, Nosofsky, and Townsend (2008) carried out an experiment using response times within systems factorial technology (e.g., Townsend & Ashby, 1983; Townsend & Nozawa, 1995) using such contrasting dimensions and showed that the results corresponded to the theoretical expectations.

In addition, the potential range of application appears very broad indeed. As a single example, we have recently carried out a static General Recognition Theory experiment (Blaha, Silbert, & Townsend, 2011). It was discovered that two fundamental dimensions of race, skin tone and physiognomy, strongly violated perceptual separability and perceptual independence, and in ways expected by certain theoretical starting points. We are excited about the possibility of extending this work to our more general purview, and the consequent development of dynamic, race-perception models.

Acknowledgment

This work was supported by AFOSR FA9550-07-1-0078.

References

Ashby, F. G. (1988). Estimating the parameters of multidimensional signal detection theory from simultaneous ratings on separate stimulus components. *Perception & Psychophysics*, *44*, 195–204.

Ashby, F. G. (1989). Stochastic general recognition theory. In D. Vickers, & P. L. Smith (Eds.), *Human information processing: measures, mechanisms, and models* (pp. 435–457). Amsterdam: Elsevier.

Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, *44*, 310–329.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33–53.

Ashby, F. G., & Maddox, W. T. (1990). Integratin information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 598–612.

Ashby, F. G., & Maddox, W. T. (1991). A response time theory of perceptual independence. In J. Doignon, & J. Falmagne (Eds.), *Mathematical psychology: current developments* (pp. 389–413). Springer-Verlag.

Ashby, F. G., & Maddox, W. T. (1992). Complex decision rules in categorization: contrasting novice and experienced performance. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 50–71.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372–400.

Ashby, F. G., & Maddox, W. T. (1994). A response time theory of separability and integrality in speeded classification. *Journal of Mathematical Psychology*, *38*, 423–466.

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, *95*, 124–150.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *91*, 154–179.

Böckenholt, U. (1992). Multivariate models of preference and choice. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 89–114). New Jersey: Lawrence Erlbaum Associates.

Blaha, L., Silbert, N., & Townsend, J. (2011). A general recognition theory study of race adaptation. Presented at the 2011 Vision Sciences Society Annual Meeting.

Busemeyer, J., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.

Cornes, K., Donnelly, N., Godwin, H., & Wenger, M. (2011). Perceptual and decisional factors influencing the detection of the Thatcher illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 645–668.

Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika*, *58*, 281–314.

Dzhafarov, E. N. (2002). Mutlidimensional Fechnarian scaling: perceptual separability. *Journal of Mathematical Psychology*, *46*, 564–582.

Ennis, D. M., & Mullen, K. (1992). A general probabilistic model for triad discrimination, preferential choice, and two-alternative identification. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 115–122). New Jersey: Lawrence Erlbaum Associates.

Farris, C., Treat, T., & Viken, R. (2010). Alcohol alters men's perceptual and decisional processing of women's sexual interest. *Journal of Abnormal Psychology*, *119*, 427–432.

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

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley and Sons, Inc.

Kadlec, H., & Townsend, J. T. (1992a). Signal detection analyses of dimensional interactions. In F. G. Ashby (Ed.), *Mutildimensional models of perception and cognition* (pp. 181–227). New Jersey: Lawrence Erlbaum Associates.

Kadlec, H., & Townsend, J. T. (1992b). Implications of marginal and conditional detection parameters for the separabilities and independence of perceptual dimensions. *Journal of Mathematical Psychology*, *36*, 325–374.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.

Lamberts, K., Brockdorff, N., & Heit, E. (2003). Feature-sampling and random-walk models of individual-stimulus recognition. *Journal of Experimental Psychology: General*, *132*, 351–378.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.

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

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

MacMillan, N. A., & Creelman, C. D. (1991). *Detection theory: a user's guide*. Cambridge: Cambridge University Press.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147–180). New Jersey: Lawrence Erlbaum Associates.

Maddox, W. T. (1995). Base-rate effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 288–301.

Maddox, W. T., & Ashby, F. G. (1996). Perceptual separability, decisional separability and the identification-speeded classification relationship. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 795–817.

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1459–1482.

Marley, A. A. J. (1992). Developing and characterizing multidimensional thurstone and luce models for identification and preference. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 299–334). New Jersey: Lawrence Erlbaum Associates.

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

Richler, J., Gauthier, I., Wenger, M., & Palmeri, T. (2008). Holistic processing of faces: perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 328–342.

Schulze, H. H., Baurichter, W., Gerling, W., & Grobe, R. (1977). Independent feature processing in the visual system: the effects of symmetry and distance of components. *Psychological Research*, *39*, 249–259.

Silbert, N.H. (2010). *Integration of phonological information in obstruent consonant identification*. Ph.D. Thesis Indiana University.

Silbert, N. H. (2012). Syllable structure and integration of voicing and manner of articulation information in labial consonant identification. *Journal of the Acoustical Society of America*, *131*.

Silbert, N.H., & Thomas, R.D. (2012). Decisional separability, model identification, and statistical inference in the general recognition theory framework. *Psychonomic Bulletin & Review* (in press). http://www.nhsilbert.net/docs/papers/Silbert_Thomas_DS_in_GRT_preprint.pdf.

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.

Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, *32*, 135–168.

Soete, G. D., & Carroll, J. D. (1992). Probabilistic multidimensional models of pairwise choice data. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 61–88). New Jersey: Lawrence Erlbaum Associates.

Thomas, R. D. (1996). Separability and independence of dimensions within the same-different judgement task. *Journal of Mathematical Psychology*, *40*, 318–341.

Thomas, R. D. (1999). Assessing sensitivity in a multidmensional space: some problems and a definition of a general *d*'. *Psychonomic Bulletin & Review*, *6*, 224–238.

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

Townsend, J. T., Hu, G. G., & Ashby, F. G. (1981). Perceptual sampling of orthogonal straight line features. *Psychological Research*, *43*, 259–275.

Townsend, J. T., Hu, G. G., & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence for positive interdependencies. *Perception & Psychophysics*, *36*, 35–49.

Townsend, J. T., Hu, G. G., & Kadlec, H. (1988). Feature sensitivity, bias, and interdependencies as a function of energy and payoff. *Perception & Psychophysics*, *43*, 575–591.

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

Townsend, J. T., & 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.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.

Wandmacher, J. (1976). Multicomponent theory of perception: feature extraction and response decision in visual identification. *Psychological Research*, *39*, 17–37.

Wickens, T. D., & Olzak, L. A. (1992). Three views of association in concurrent detection ratings. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 229–252). New Jersey: Lawrence Erlbaum Associates.