Selective Influence and Classificatory Separability (Perceptual Separability) in Perception and Cognition: Similarities, Distinctions, and Synthesis

5

James T. Townsend, *Yanjun Liu*, *Ru Zhang*
Indiana University Bloomington, Bloomington, IN, United States

In our opinion, psychology and its close cousin (and partial offspring), cognitive science, are properly viewed as black-box sciences where for the most part, the inner workings must be divined from the orderliness of input–output relationships. Even today, physiology can offer, at best, relatively modest resources to this end, though we hope for massive improvements in the marriage of physiological recordings and behavioral technologies as the future unfolds.

One fundamental theme in psychology concerns whether a change in one entity causes an alteration in another entity or not. Unsurprisingly, the issue of invariance vs. change is fundamental in psychological science and plays out in many forms. Of course, orderly dependence is the obverse of independence or invariance. We earlier have defined the very general *principle of correspondent change* as the correlative structure and function in a theory with concomitant change, or invariances, in nature,[1] such as parameter invariance across certain variations of experimental circumstance (Townsend & Ashby, 1983, Chapter 15, p. 481). Absent this precept, mathematical applications to science are little more than descriptive exercises at best, or useless, scientifically vacant meanderings at worst. The present concepts are special cases of this foundational notion.

Of course, where the element of chance is ubiquitous, as within psychology, dependence and independence have to be interpreted in a probabilistic framework. This aspect, by itself, forms no serious challenge since dependence, or no, has long been a hallmark of probability theory.[2,3]

Two types of functional dependence that have played a central role in our own theories and methodologies are *selective influence* and *perceptual separability*. There have been hundreds of papers utilizing these concepts.[4] It would be especially agreeable if the investigator were able to directly observe samples of the random variables themselves. With regard to selective influence, the operative observable random variable has usually been response time (*RT*) and with separability, usually confusion frequencies. For instance, if we could observe the actual processing times of the processes in a serial system, we would probably not require the more removed, but as it turns

Systems Factorial Technology, DOI: 10.1016/B978-0-12-804315-8.00007-0

out, essential, statistics on response times. Similarly, in general recognition theory, we can typically only observe probabilities of recognition and confusion rather than the underlying percepts themselves.

More rigorous and deeper characterizations will be offered subsequently, although we strive for communicability at the possible expense of a stark logico-deductive approach. To initiate the discussion, selective influence lies at the heart of the popular *additive factors method* (Sternberg, 1969), a methodology which can support or disconfirm serial processing. It also underpins the subsequent *systems factorial technology* (SFT), which permits testing of parallel and more exotic mental architectures (e.g., Townsend & Nozawa, 1995; Townsend & Wenger, 2004; see Algom, Eidels, Hawkins, Jefferson, & Townsend, 2015, for history and general discussion; of course, this volume contains a treasure trove of this theory, methodology, and applications), stopping rules, capacity, and independence. Its power derives from its mathematical rigor but also that it provides tests that are nonparametric as well as distribution free. These characteristics are almost unique in the annals of the study of perceptual, cognitive, and action systems. Informally, selective influence of an experimental factor means that exactly one of the two or more psychological processes are affected by that factor. Selective influence is essential to uncovering the architecture and stopping-rule of the attendant cognitive processes.

Now let us dig deeper into the concept of selective influence with an experimental example. This study by Wenger and Townsend (2001) used the *double factorial paradigm*, invented by Townsend and Nozawa (1995; the present volume celebrates the 20th anniversary of that publication). One factor is simply the presence vs. absence of each target (typically two) whereas the other factor is really a set of two factors meant to speed up or slow down processing of one or the other, or both, targets.

The Wenger and Townsend double factorial paradigm investigated detection of presence vs. absence of two target features, eyes and mouth, either in standard position in a normal face stimulus or as location-disorganized features. For our example, we now consider the normal face. The experimental factors consisted of the degree of visual clarity of the separate features.

Taking combinations of different levels of the clarity factors on trials when both targets were present, they created a double-factorial design of 4 stimuli: high-clarity eyes^high-clarity mouth (hh), high-clarity eyes^low-clarity mouth (hl), low-clarity eyes^high-clarity mouth (lh), low-clarity eyes^low-clarity mouth (ll). It was expected, and confirmed, that a higher level of salience of either feature should lead to faster perception of that feature.

The early usage of selective influence was only concerned with demonstrating influence at the mean *RT* level. Townsend and colleagues (Townsend, 1984, 1990a; Townsend & Schweickert, 1989) demonstrated that in order to empirically distinguish parallel from serial systems, it was necessary for the selective influence operate at a more fundamental level. Recall that the *cumulative distribution function* for *RT*s yields the probability, for an arbitrary time t, that the response was made faster than that duration. Likewise, the so-called *survivor function* is just 1 minus the cumulative distribution function and gives the likelihood that the response took longer than any given t. The "survivor" part of the term comes from its use in actuarial statistics in

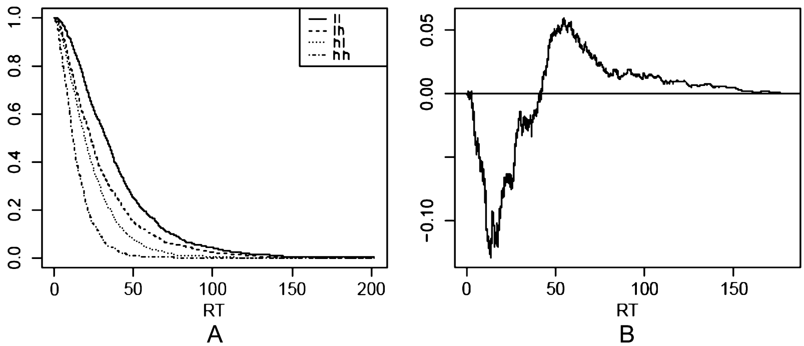

Figure 5.1 (A) Distributional ordering holds for survival functions of simulated response time of ll, lh, hl, and hh. This indicates the fulfillment of selective influence. (B) Estimated SIC function from the simulated data. The SIC pattern indicates the two processes are arranged in a serial exhaustive manner.

calculating the probability that a person, machine, etc., survives beyond any certain point of time, t. Then, a sufficiently strong level of influence, even at the level of mean *RT*s, is that the factors order the *RT* cumulative distribution functions, or equivalently, the survivor functions (e.g., Townsend, 1990a). Significantly, when selective influence fails, then at least at the level of mean *RT*s, virtually any pattern of interactions can occur (Townsend & Thomas, 1994).

Ensuing developments showed that the same level of selective influence as that testing mean *RT*s, made much more distinctive predictions for the distribution functions themselves (Townsend & Nozawa, 1995). Thus, once the selective influence holds at the level of distribution ordering, SFT would predict unique patterns of *survival interaction contrast* (SIC) that corresponds to a certain combination of architecture (serial, parallel, or coactive) and stopping rule (self-termination or exhaustive). Figs. 5.1A and 5.1B are examples of the fulfillment of selective influence at the distributional ordering of survival functions and the estimated SIC from simulated data.

These unique patterns therefore help to unveil the human's cognitive systems and their dynamics (e.g., as an impressive application, see Wenger & Ingvalson, 2003). More technical details and discussion of selective influence will unfold later on in the chapter.

Perceptual separability has been more confined to perception per se at least in literature that employs precisely that name and has not received quite as much theoretical analysis as selective influence. It is nevertheless an exceedingly important topic in human information processing and its identifiability in data continues to be investigated (Garner & Morton, 1969; Ashby & Townsend, 1986; Kadlec & Townsend, 1992; Maddox, 1992; Ashby & Soto, 2015). Again informally, its defining characteristic is that, given two dimensions or features, the percept of one psychological dimension is invariant across stimulus changes in the other (see also Silbert & Thomas, 2013; Soto, Vucovich, Musgrave, & Ashby, 2015).

For a starting example, consider the perceptual dimensions of loudness and pitch. Then, perceptual separability stipulates that the probability of perceiving one level of a dimension is constant across different levels of the other dimension. Similarly,

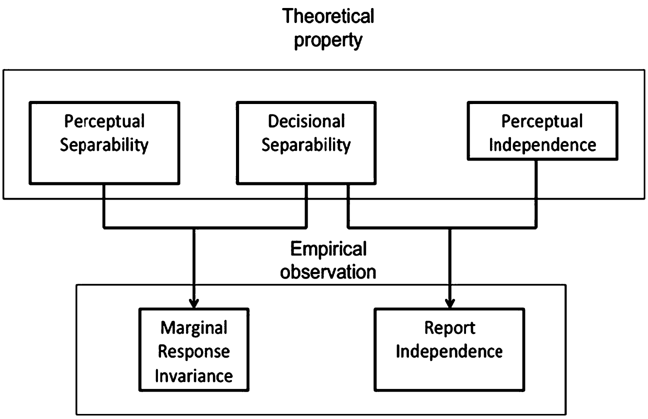

Figure 5.2 Theoretical structure of General Recognition Theory.

in the example from above of selective influence on the facial features of mouth and eyes, perceptual separability demands that the perception of eyes, say, be invariant across the presence or absence of the other. If two distinct sets of eyes or mouths were employed as stimuli then again, the ability to say, perceive which mouth was present, would not be a function of the level of the eyes.

Although observing the process of perception is impossible, *general recognition theory* (GRT, Ashby & Townsend, 1986; Ashby & Soto, 2015) provides us with a mathematical approach that could induce the perceptual separability and other underlying cognitive processes from empirical observations of response frequencies. The theoretical structure of GRT is summarized in Fig. 5.2 and the detailed description of the theory can be found in Ashby and Townsend (1986), Kadlec and Townsend (1992) and Soto et al. (2015).

GRT predicts that if we observe *marginal response invariance* such that the distribution of response to one level of dimension is invariant across different levels of the alternative dimensions, then we could induct the fulfillment of perceptual separability. Thus, if pitch and loudness perception are perceptually separable, then manipulations of either stimulus pitch or loudness should only affect the respective perceptual dimension. It is obvious that this concept of invariance readily extends beyond perception per se.

Interestingly, in addition to the fact that both these vital concepts have something to do with cognitive invariances, it turns out that they share certain profound characteristics. Selective influence and perceptual separability are similar concepts to each other in the sense that both depict the relation that the percept of one dimension is invariant across manipulation of the orthogonal dimensions. Selective influence was originally response time-oriented; whereas perceptual separability was originally response classification.

We will later come to see that although selective influence has received more theoretical analysis, some of the same questions and characteristics also pertain to questions about separability. In addition, we will subsequently contemplate whether either classificatory separability or selective influence could be enlisted in the other's goals.

To our knowledge, there has never been an investigation aimed at explicating their similarities and, if there are any except for linguistics, substantive differences. In the present enquiry, we think it is important to include not only what appear to be deep-seated mathematical or semantic distinctions or resemblances, but also the purpose and intent of the users. Typical conventions will also be grist for discussion.

As pointed out above, the term "selective influence" implies manipulation of something with psychological relevance in order to produce evidence from an observable variable that is intended to reveal some type of psychological structure or function. Perceptual separability, on the other hand, refers to a dimensional or featural percept remaining invariant with a stimulus change on a distinct dimension or feature.

However, the epistemological implications of this part of the discussion go far beyond perception as will become more apparent below. We thereby coin the new term *classificatory separability* to substitute for "perceptual separability." In fact, the following characterizations intend to be inclusive of earlier usage, but attempt to expand their generality and potential scope of applicability. This tack is especially helpful when discussing the scientific purpose of the concepts. References will be given to literature which makes available the mathematical underpinnings.

Selective Influence

The basics tenets of selective influence are:

I. The purpose of selective influence. The experimenter manipulates one or more variables (typically at least two) with the design of separately affecting distinct psychological processes.

II. The selective influence variable may be external or internal to the observer: (A) It is external if applied to an external sense organ, for instance, the eyes, ears, skin, etc.; (B) It is internal if applied to an internal organ, sensory or not. For instance, an ingested pharmaceutical might induce a migraine headache (sensory) vs. causing slowed gastric motility without sensation (non-sensory).

III. The selective influence variable may be qualitative or quantitative, though the quantitative instances, if not universal, were the first employed and still appear to form the most potent in reaching the aims of the researcher.

IV. If quantitative, the variables could be unidimensional or multidimensional and the dimensionality may or may not be expressly defined. Expressly defined dimensionality could be features of an object such as brightness, shape; whereas the implicitly dimension could be psychological dimensions.

V. The experimenter specifies not only the putative selective influence variable but also one or more dependent variables, which have been shown to respond differentially to the stipulated selective influence variables.

VI. The selective influence variable may be part of the stimulus with which the observer is presented, or not.

VII. If part of the stimulus, the selective influence variable may be an aspect or dimension which directly relates to the observer's task or it might be unrelated to the task.

Most of the foregoing may be immediately evident but (VI) and (VII) might be unpacked a bit. An example of a case when the selective influence variable is not part of the stimulus is the following: Suppose it has been hypothesized that the motor system of the observer will act more slowly when the ambient temperature is turned up higher than normal but that this operation does not affect early perceptual processing times. This manipulation is not an inherent part of the observer's task. A case of a selective influence variable which is an integral part of the stimulus would be the brightness of a visual stimulus which the observer must perceive.

With regard to (VII), suppose now that light intensity as a selective influence variable is used where the actual task of the observer is to identify an English letter from the usual 26 letter alphabet. The light intensity itself is not part of the task. However, if the similarity of the stimulus letters is manipulated, this facet may be considered to be an intimate particulate of the basic task. A reviewer points out that the aforementioned ambient temperature and stimulus brightness might be assessed with regard to the architecture of their processing.

The prior discussion holds whether interpreted probabilistically or deterministically. Now assume, as in most psychological milieus, we are dealing with probabilities.

Definition 1. A *Psychological System* is a collection of connected entities called *processes*.

Every process is usually taken to interact with at least one other process to avoid triviality.

Definition 2. A *Set of Experimental Factors* is a collection of entities that can affect one or more of the subprocesses under study.

Definition 3. *Process* is a proper part of a psychological system, which performs a psychological function. It is assumed to have some set of inputs and a set of outputs and usually is expected to expend time to do its job.

Definition 4. *Influence* by an experimental factor on a process implies that the probability distribution on some dependent variable affected by that process is altered under change of the factor values.

If an experimental factor affects a single process, its influence is said to be *selective*.[5] The original definition from Sternberg (1969) did not make entirely clear what might or might not satisfy selective influence, and the emphasis was on mean *RT* and thus the means of the pertinent random times. Suppose that processing times T_x and T_y are bivariate normally distributed, there are three properties to characterize T_x (T_y): the mean of T_x (T_y), the variance of T_x (T_y), and the covariance between T_x and T_y. If the distribution of T_x is affected by a variable X while the distribution of T_y is affected by a distinct variable Y and T_x and T_y were stochastically independent, then

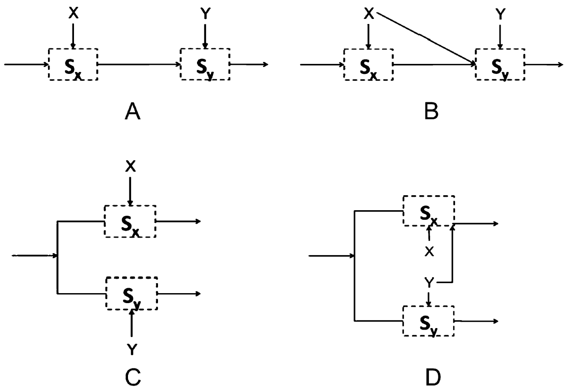

Figure 5.3 (A) Selective influence fulfilled in a serial system with experimental factors: X, Y. (B) A direct non-selective influence in a serial system with experimental factors X and Y. (C) Selective influence fulfilled in a parallel system with experimental factors X and Y. (D) A direct non-selective influence in a parallel system with experimental factors X and Y.

selective influence could be claimed on the distributional level. If T_x and T_y are interdependent, one cannot claim selective influences without inquiring the cause of the interdependence. Having the mean of T_x depend on X and the mean of T_y depend on Y, the independence between the means and the stochastic dependence of T_x and T_y implies that the stochastic dependence of T_x and T_y depends on some variable that is independent of X and Y. Hence manipulating $Y(X)$ influences neither the mean of $T_x(T_y)$ nor the covariance between T_x and T_y. Then selective influences at the level of mean RTs can be claimed. Attempts were made subsequently to deepen the definition as in the following.

Definition 5. (E.g., Townsend, 1984; Townsend & Schweickert, 1989.) Consider two processes S_x and S_y with processing times T_x and T_y. Selective influence holds if and only if there exist two experimental variables, X and Y, such that T_x is affected only by X and T_y is affected only by Y.

This definition does move the emphasis on means to that of the underlying random time variables, but it leaves fuzzy just what role stochastic dependencies might play. This important aspect will be further developed below.

Townsend and colleagues (Ashby & Townsend, 1980; Townsend, 1990b; Townsend & Ashby, 1978; Townsend & Schweickert, 1989) have shown that how, and to what depth, an experimental factor affects the observable statistic is strongly linked to how much of an operating system can be identified from that statistic. Figs. 5.3A through 5.3D exhibit the fundamental serial and parallel architectures, which either satisfy, or not satisfy, selective influence. *Direct non-selective influence* is said to occur if an experimental factor has a direct impact on the "wrong" process, for example, factor X affects T_y. A more subtle type of failure will be discussed later.

The way in which a factor effect operates probabilistically can simply be defined through a deformation of a probability distribution, which leads to the next pivotal definition.

Definition 6. An experimental factor influences a processing time distribution at the level of *stochastic dominance*, if it orders the distribution so that $F(t; X_2) = P(T \leq t; X_2) > F(t; X_1) = P(T \leq t; X_1)$ if and only if the experimental factor satisfies $X_2 > X_1$.

Stochastic dominance is a fairly weak type of dominance in a hierarchy of potential echelons of influence (Townsend, 1990a; Townsend & Ashby, 1978). However, influence only at the even more coarse level of mean processing times, an example of an 'ineffective' influence in most studies of response times, is too weak to prove the important theorems pertaining to testing various architectures and stopping rules (e.g., Townsend, 1984; Townsend & Ashby, 1983).

For instance, if the researcher is attempting to assess whether visual identification of separate letters takes place in serial or parallel, then one may manipulate brightness and determine whether or not the response time cumulative distribution function on an individual letter is always larger at any arbitrary time for the brighter stimulus. When Schweickert (1978) began to expand the kinds of architecture that could be identified by factorial manipulations, he assumed that psychological process durations were affected by adding a time increment (or decrement) to each process time. The main results at that time were confined to deterministic systems, that is, systems for which the constituent process durations were constant, although Schweickert presented certain RT bounds for stochastic situations (e.g., Schweickert, 1982).

Taking this concept to fully stochastic systems, it had long been known that if a positive and independent increment is added to a random variable, stochastic dominance occurs. However, this situation is a special case of the latter. Townsend and Schweickert (1989) showed that if one gives up the idea of independence of the increment, then stochastic dominance and adding a positive random increment to the random variable are mathematically equivalent.[6,7]

It is important before proceeding to re-emphasize that the simple satisfaction of selective influence will not automatically lead to, for instance, identification of parallel vs. serial processing. In fact, even if T_x is affected only by X and T_y is affected only by Y with attendant stochastic independence, tests of parallel vs. serial processing at the level of mean *RT*s will not be feasible (e.g., Townsend & Ashby, 1983; Townsend, 1984). As an apt example, suppose that the variance of T_x is affected by X but the mean of T_y is influenced by Y. It is not clear how to even necessarily detect these effects, much less how to use them to determine aspects such as architecture.

As another example, consider that the prediction by a parallel system with a minimum time stopping rule is that the means, as functions of the experimental factors, should exhibit a certain interaction (in fact, positive in this case as opposed to zero in the serial model). If the experimental factors order only the means of the individual channels, then this interaction could be anything.

Figs. 5.4 and 5.5 illustrate these concepts. First, Fig. 5.4A shows dominance at the mean level. Since, it appears that one distribution is simply a shift of the other, one expects dominance at the distributional level to hold as well (Townsend, 1990a), which is illustrated in Fig. 5.4B. Fig. 5.4C then indicates a valid SIC signature for parallel processing with an exhaustive stopping rule, when dominance at the level of Fig. 5.4B is true.

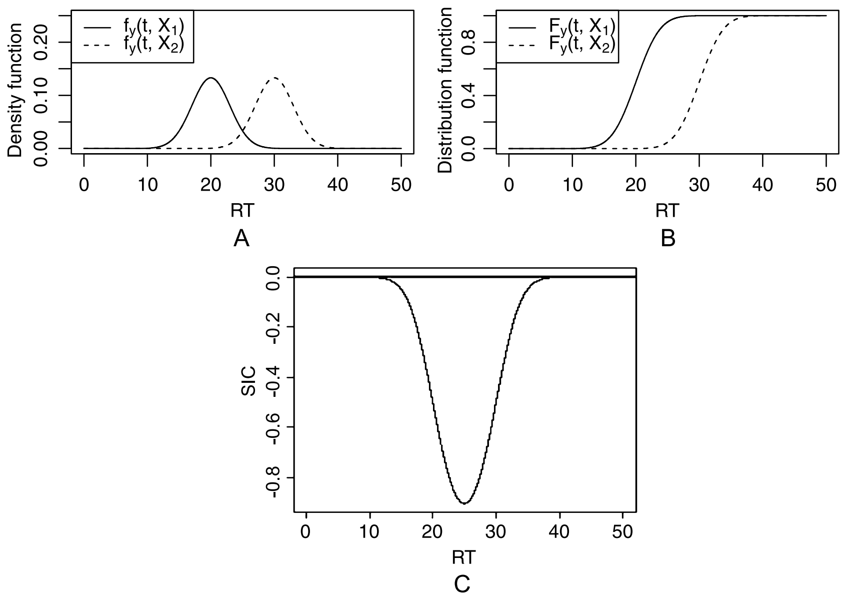

Figure 5.4 (A) Stochastic dominance holds at the mean level. (B) Stochastic dominance holds at the distributional level. (C) Survival interaction contrast of parallel exhaustive processing with stochastic dominance holding at the distributional level.

On the other hand, Figs. 5.5A, 5.5B, and 5.5C reveal what can go wrong if mean dominance occurs but distributional dominance fails. Observe in Fig. 5.5C that an invalid SIC signature results.

Dzhafarov and colleagues (e.g., Dzhafarov & Kujala, 2012; see also, footnote 5) have performed a series of mathematical analyses on selective influence at the distributional level. The most important specification from this sector, for present purposes, was that the germane process random variables could be probabilistically dependent on a third random variable which was itself totally independent of the experimental factors (e.g., Dzhafarov, 2003). A simplified special case of his definition, but which will suffice for our purposes, follows.

Definition 7. (Dzhafarov, 2003.) Selective influence holds for random times T_x and T_y if and only if T_x is a function of X and C whereas T_y is a function of Y and C. X and Y are distinct experimental variables while C is a random variable independent of X and Y.

It can be readily shown that if $C = c$ a particular value, T_x and T_y are conditionally independent (as shown in Fig. 5.6). Furthermore and more importantly, all the theorems of Townsend and Nozawa (1995) that permit identification of process characteristics such as architecture and stopping rule then go through immediately (e.g., Dzhafarov, 2003). This issue will be revisited below.

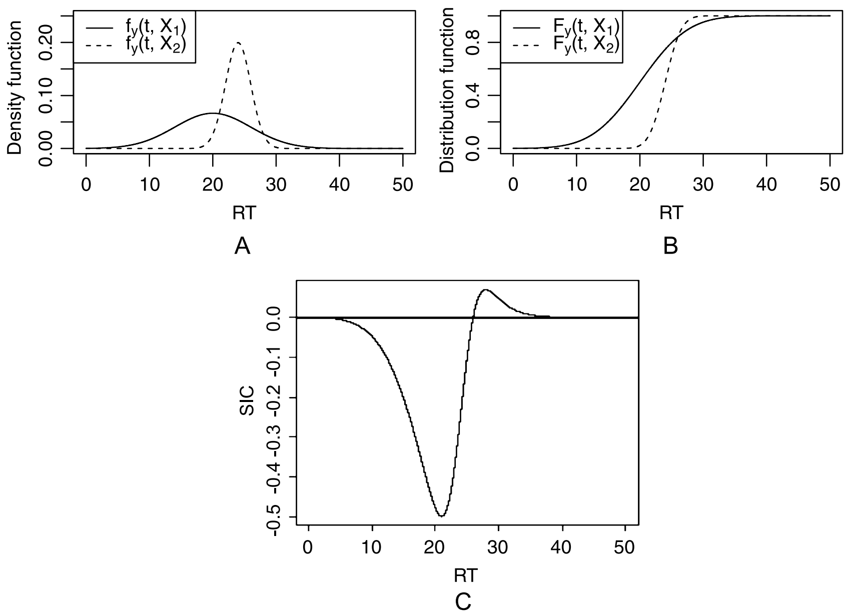

Figure 5.5 (A) Stochastic dominance holds at mean-level. (B) Stochastic dominance fails at the distributional level. (C) SIC with absence of stochastic dominance at the distributional level. The shape of this SIC function does not follow any typical four SIC functions defined in Townsend and Nozawa (1995). But it is similar to the characteristic SIC pattern of serial exhaustive processing.

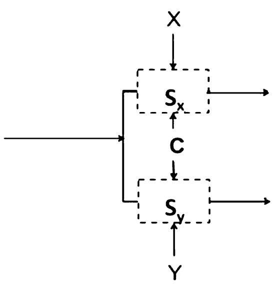

Figure 5.6 A parallel system with experimental factors: X and Y. X, Y, and C are independent of each other. For this case, conditional independence (conditioning on C) holds.

Classificatory Separability

To make sense of our generalization of perceptual separability, we need the concept of a collection of entities that belong to a psychological family of some sort, some type of resemblance as it were. This kind of invariance or not, as observed earlier, has been of interest in psychology back to its philosophical roots. In psychophysics, it has been of concern whether, for example, the perception of pitch changes when loudness is changed through manipulation of the sound intensity and vice versa. It has long been known that neither is invariant as the other dimension of the stimulus is altered.

A full panoply of types of independence, when response frequencies are analyzed, is presented in Ashby and Townsend (1986). Related early statements can be found in Kadlec and Townsend (1992) and Maddox (1992).

Psychological dimensions such as expression and age of a face form another example. For instance, an informal observation suggests that a middle aged face appears older in wearing a dour expression and younger with a happy expression. This particular example is probably asymmetric since it seems less likely that the degree of, say, happiness or sadness might be perceived as more extreme in old vs. young faces. In a deterministic environment, where, for example, faces are completely visible, without noise, the invariance of expression with age or its failure, are obvious (e.g., Townsend & Spencer-Smith, 2004). But what about viewing a face through fog or in a bad photograph? We require interpretations in terms of probability distributions again.

Then, we define classificatory separability as the invariance of a particular member of that class when experienced consciously when another class is changed. Thus, in more rigorous terms:

The basic tenets of classificatory separability are:

I. Psychological class A is classificatory separable from B if the identity of a value from A is unaffected by a change in class B.

II. Classificatory separability or its failure can occur: (i) due to manipulations by an experimenter, or (ii) can occur autonomously.

III. If II(1), the experimenter stipulates one or more dependent variables which refer to the processing of A and B and that will reflect in some way whether or not invariance eventuates.

Let A_i represent the levels of class A, and B_j those of class B with $i,\ j\ =\ 1, 2$ within the context of a so-called *complete identification paradigm*, that is, one where all four combinations of class values appear across the trials of the experiment. Also, let $g_{ij}(a, b)$ be the joint distribution of psychological values a and b, where a is associated with A and b with B, for stimulus compound $A_i B_j$.

The original definition of perceptual separability in Ashby and Townsend (1986) was as follows, with that term replaced by "classificatory separable". It is also sketched in Fig. 5.7.

Definition 8. Psychological class A is classificatory separable across levels of B for a given level A_i if and only if

$$g_{i1}(a, b) = g_{i2}(a, b).$$

Similarly, psychological class B is classificatory separable across levels of A for a given level B_j if and only if

$$g_{1j}(a, b) = g_{2j}(a, b).$$

Observe that separability could hold for one level of a class but not the other and/or for one class but not the other. As noted above, such a situation might obtain in that perceived age might be a function of expression but not vice versa. The marginal

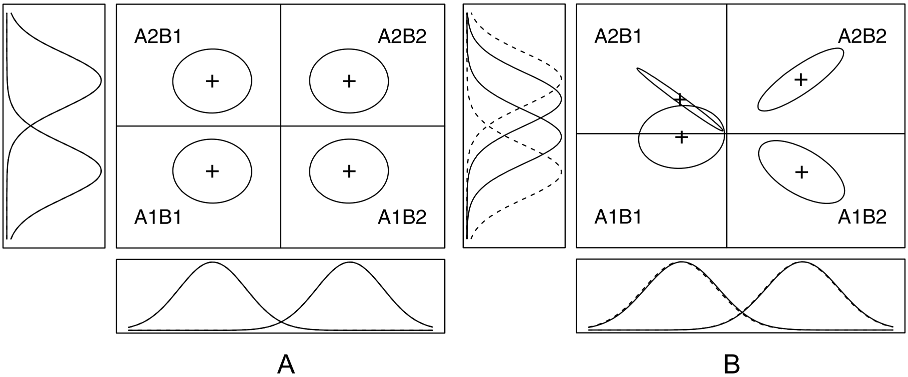

Figure 5.7 (A) The four ellipses are the contours of the joint distributions of responses to the two dimensions. The marginal distributions are plotted on the bottom and on the left. Perceptual separability holds for both dimensions: the marginal distributions of each dimension are invariant across the change of the other stimulus dimension. (B) Perceptual separability fails for dimension A as marginal distributions at each level of dimension A is different across the change of dimension B. However, the separability still holds for dimension B.

distribution on perceived age of an individual would be shifted higher when paired with the dour expression as opposed to that paired with a happy expression.

Finally, it is important to grasp that those interactions among percepts or other categories could occur, just as in selective influence any one of a hierarchy of relations might obtain (Townsend, 1990a). With that fact in mind, we next delve into an essential quantitative correlate of both our major concepts. That notion is the marginal distribution of the central joint distributions, processing times in the case of selective influence and perceptual or cognitive observation random variables in the case of separability.

The Pivotal Notion and Role of Marginal Selective Influence

Marginal Selective Influence in RT Studies

Since selective influence has received more theoretical analysis than separability, over the past two or three decades, we first take up selective influence with respect to *marginal selective influence*.

Townsend and colleagues (e.g., Townsend & Ashby, 1983; Townsend, 1984; Townsend & Thomas, 1993) allowed for probabilistic interactions of the processes as long as the marginal distributions remained unaffected by the "wrong" experimental factors. For instance, in the serial case, processes might be *across-stage independent*, or *dependent* meaning that the second process in a serial chain could depend stochastically (e.g., be on average faster or slower) on the first, depending on how long the first

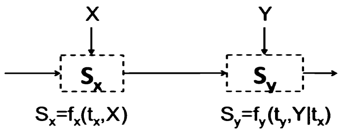

Figure 5.8 A serial system with the potential for indirect non-selective influence due to the dependency of density function of the second stage on the time taken by the first stage.

process took to complete. Let T_x be the random time associated with the first stage in a serial system and T_y that for the second stage. Let t_x and t_y be the specific values for T_x and T_y, respectively. X and Y are experimental variables, which the experimenter hopes would separately affect the first and second stages, respectively.

Definition 9. I. Across-stage independence occurs if the density functions satisfy $f_y(t_y, Y|t_x) = f_y(t_y, Y)$. II. Across-stage dependence occurs if and only there exist values of t_x such that $f_y(t_y, Y|t_x) \neq f_y(t_y, Y)$.

Suppose the two stages were across stage independent, then $f_y(t_y, Y|t_x) = f_y(t_y, Y)$. And assume there is no direct non-selective influence so that $f_y(t_y, Y, X|t_x) = f_y(t_y, Y|t_x)$. Then as a direct consequence of across stage independence $f_y(t_y, Y|t_x) = f_y(t_y, Y)$, selective influence follows perforce.

Next, what about when across stage independence fails, that is, $f_y(t_y, Y|t_x) \neq f_y(t_y, Y)$, and the left hand side is a non-trivial function of the previous stage time, t_x? We can consider the density on T_y, written as a random function, depending on the first stage time t_x, $f(T_y, Y|T_x = t_x)$. When across stage dependence occurs, as indicated, then the (random) distribution of $f(t_y, Y)$, conditioned on $T_x = t_x$ will be a non-trivial function of t_x: $E_{ty}[f(T_y, Y|T_x = t_x)] = f(t_y, Y|T_x = t_x)$ is clearly a function of t_x and not, ordinarily, equal to its marginal distribution. In fact, the latter will usually be a function of the wrong experimental variable X. We may find it by taking the expectation of $f(T_y, Y|T_x = t_x, X)$ over $T_x = t_x$:

$$E_{tx}\left\{E_{ty}\left[f(T_y, Y|T_x, X)\right]\right\} = \int E_{ty}\left[f(T_y, Y|t_x, X)\right] f(t_x, X)dt_x = f(t_y, Y, X).$$

The integral, of course, is taken over $0 < t_x < \infty$. Observe that the final marginal density for t_y is a non-trivial function of the "wrong" factor X!

Definition 10. *Potential indirect non-selective influence* occurs when there is across stage dependence.

Fig. 5.8 indicates the incidence of indirect non-selective influence in a serial system. Nonetheless, if somehow the dependence on X disappears on marginalization, then selective influence still is in force. This will not ordinarily occur for arbitrary conditional distributions $f(T_y, Y|T_x = t_x)$, but can under certain circumstances. In that event we have an instance of Definition 11.

Definition 11. *Marginal selective influence* holds if and only if $E_{tx}\{E_{ty}[[f_y(T_y, Y|T_x)]] = f_y(t_y, Y)$.

Just to firmly instantiate this last principle, suppose $f_y(t_y, Y|t_x) \neq f_y(t_y, Y)$, that is, across stage independence does not hold and potential indirect non-selective influence (Definition 10) does. Then, it could then be, and probably will be, that the marginal distribution on T_y, after integrating over T_x, will, indirectly, be a function of X as detailed just before Definition 10. It is an instance of indirect non-selective influence. We formalize this concept in Definition 12.

Definition 12. If X does not directly affect T_y yet $E_{tx}[f_{ty}(T_y, Y)|T_x = t_x, X] = f_y(t_y, Y, X) \neq f_y(t_y, Y)$, then marginal selective influence fails and indirect non-selective influence is said to occur.

At the time (e.g., Townsend & Nozawa, 1995; Townsend & Schweickert, 1989) it was not investigated as to whether, when and how the "marginalizing out" might occur, when potential indirect non-selective influence was present, due to dependence across stages. But, note that if Dzhafarov's definition of selective influence is satisfied (e.g., Dzhafarov, 2003; see Definition 7 above) then marginal selective influence will indeed be satisfied. Thus, even with the dependence of two random variables say T_x and T_y on a third random variable, C in addition to selective factors, respectively X and Y, marginal selectivity could, in principle, hold and, further including the assumption of distribution ordering mentioned above, permits the proofs of architecture identification to go through unimpeded. However, if marginal selectivity does not hold, then havoc even in mean *RT* predictions erupts (Townsend & Thomas, 1994).

Further, as Dzhafarov and Kujala (2010) demonstrate, even a stronger version of marginal selectivity does not imply Dzhafarov's definition of selective influence in the presence of certain kinds of stochastic dependencies. They have produced a straightforward counter example, based on discrete probability distributions, in which the variables do satisfy even a somewhat more stringent version of marginal selectivity, yet fail to actually fulfill the condition of selective influence (Kujala & Dzhafarov, 2010). This counter-example may be thought of as an instance of indirect non-selective influence. Thus, at least within the class of discrete distributions, marginal selectivity is a necessary but not sufficient condition for the Dzhafarov (2003) definition of selective influence. Nonetheless, it is not at all clear whether marginal selectivity might be sufficient when in the presence of, say, continuous probability density functions that are continuous in their experimental variables. Moreover, we suspect that a supplementation of such densities with the strong condition of distribution ordering by the factors may be enough to force sufficiency. We are presently exploring this branch of the field.

At this point in time, there exists a test that is a necessary and sufficient condition for selective influences on finite discrete variables (Dzhafarov & Kujala, 2010). This is a test that can be performed without having to assume the distribution ordering. It is called the *Linear Feasibility Test*. In a complete identification paradigm if the test is passed, or in other words, selective influence on discrete variables is established, if one can find four jointly distributed variables $(H_{X1}, H_{X2}, H_{Y1}, H_{Y2})$, each corresponding to a level of a factor, satisfy the following constraints:

$$\sum_{ln} \Pr(H_{X1} = k \wedge H_{X2} = l \wedge H_{Y1} = m \wedge H_{Y2} = n)$$
$$= \Pr(T_{X1} = k \wedge T_{Y1} = m \mid X_1 Y_1),$$
$$\sum_{lm} \Pr(H_{X1} = k \wedge H_{X2} = l \wedge H_{Y1} = m \wedge H_{Y2} = n)$$
$$= \Pr(T_{X1} = k \wedge T_{Y2} = m \mid X_1 Y_2),$$
$$\sum_{kn} \Pr(H_{X1} = k \wedge H_{X2} = l \wedge H_{Y1} = m \wedge H_{Y2} = n)$$
$$= \Pr(T_{X2} = k \wedge T_{Y1} = m \mid X_2 Y_1),$$
$$\sum_{km} \Pr(H_{X1} = k \wedge H_{X2} = l \wedge H_{Y1} = m \wedge H_{Y2} = n)$$
$$= \Pr(T_{X2} = k \wedge T_{Y2} = m \mid X_2 Y_2),$$

and

$$1 \geq \Pr(H_{X1} = k \wedge H_{X2} = l \wedge H_{Y1} = m \wedge H_{Y2} = n) \geq 0,$$

where k, l, m, n are values of $T_{X1}, T_{X2}, T_{Y1}, T_{Y2}$.

There are some necessary conditions for selective influences. One is called the *distance test* (Kujala & Dzhafarov, 2008). An alternative test that can be used when one has access to covariance matrices is the *cosphericity test* (Kujala & Dzhafarov, 2008). We observe that if the distributions are multivariate Gaussian, then the following condition holds under transformations to standard form, that is, with means equal to 0 and variances equal to 1. Then, letting $r(i, j)$ be the covariances (= correlations here) for each pair of bivariate distributed variables, we have

Definition 13. (Kujala & Dzhafarov, 2008.) Correlations $r(i, j)$, $i, j = 1, 2$ satisfy *cosphericity* if and only if

$$\left| r(1,1) \cdot r(2,1) - r(1,2) \cdot r(2,2) \right| \leq \left[\left(1 - r(1,1)^2\right) \cdot \left(1 - r(2,1)^2\right) \right]^{1/2} + \left[\left(1 - r(1,2)^2\right) \cdot \left(1 - r(2,2)^2\right) \right]^{1/2}.$$

Moreover, if the researcher is willing to relinquish some generality, then by assumption of bivariate Gaussian distributions, the relevant condition above becomes necessary as well as sufficient. If, in addition, the pertinent random variables are directly observable, Kujala and Dzhafarov (2008) have adduced another test. Since it is virtually never true that the pertinent random variables are directly observable in either typical *RT* or accuracy paradigms, we do not consider it here.

As intimated earlier, a taxonomy of distributional correlations is lacking, for example, based on across-stage dependencies in serial processes that yet permit marginal selectivity. Nonetheless, it is quite important to appreciate that marginal selectivity at the level of distribution ordering suffices for a proof of the fundamental theorems of

Townsend and Nozawa (1995) on the identification of mental architectures and stopping rules. These and related predictions also appear in a number of chapters in this volume.

Marginal Selective Influence in Response Frequency Designs

Strikingly, we can immediately espy from Definition 8 that classificatory separability is equivalent to marginal selectivity in the distribution on A or B, as in Definition 11: When an experimenter changes the value of a stimulus dimension, say auditory frequency and thereby alters the perceived pitch, if the sense of loudness were unaffected at the level of the marginal distribution, then loudness would be perceptually separable from that of pitch. Since under noisy or low-energy conditions, pitch will be stochastic, the thinking was (e.g., Ashby & Townsend, 1986) that we can only ask that its distribution be unaffected.

A number of considerations long embedded in discussions of architecture and its related topics in the theoretical response time literature have been absent from those associated with accuracy and general recognition theory simply because they didn't arise naturally. Thus, on the one hand, notions such as the conditional dependencies within Definition 12 could exist, but were not treated because they failed to arise within some context analogous to, say, serial vs. parallel processing. And, the static form of the original general recognition theory and the traditional concepts associated within psychometrics, signal detection theory and multidimensional scaling, undoubtedly encouraged researchers to disassociate correlation structures from the values of the distributional means. Conversely, the milieu within the field of response time studies was and is, immediately dynamic and thus promotes such conceptions as "state of processing or activation" as intimately related to the "states of processing" of other psychological systems or stages.

The next step forward is to provide a formulation within which the concepts of selective influence and classificatory separability are special cases. But, this generality is, in fact, inherent in the definition provided by Dzhafarov (2003), recalled in our Definition 7.

A Synthesis of Classificatory Separability and Selective Influence

Both Selective Influence and Classificatory Separability Involve Psychological Processes

Earlier parts of this study have emphasized that classificatory separability and selective influence differ in their genesis, purpose, and usage. Nonetheless, the reader may well have already detected, particularly from the abstract definitions of selective influence offered above, given in terms of random entities, such as random variables, that they can be brought together in a tidy fashion.

Suppose the psychological value of a class on a trial, for instance, a percept of a psychological dimension, say A, is associated with a *process* as defined in Definition 3. And, a change in its value is associated with a value-changing *factor*, say X and similarly for dimension B and factor Y. Then, all the aspects of selective influence can be applied to perceptual separability and an extension to arbitrary classification issues is evident.

Within this new setting, we can also discern that just as selective influence can occur in different ways (e.g., simply any change to a distribution vs. some ordering associated with the distribution), so could the degree of invariance in classificatory separability. So, an alteration on dimension A could affect any aspect of the distribution such as the mean or it might order the cumulative distribution functions, hazard functions or some other statistic. For instance, with dimensions such as pitch, we won't expect a change to necessarily order the cumulative distribution functions.

Less frequently considered, even in the selective influence literature, is the possibility of failure of invariance at distinct levels. For instance, if for the distribution functions, $F_y(t_y, Y|t_{x2})\ F_y(t_y, Y|t_{x1})$ when $t_{x2} > t_{x1}$ and $F_x(t_x, X_2) > F_x(t_x, X_1)$ if $X_2 > X_1$ then perforce $E_{Tx}\{E_{Ty}[f_y(T_y, Y|T_x, X)]\} \neq f_y(t_y|Y)$ and even more intriguingly, $E_{Tx}\{E_{Ty}[f_y(T_y, Y|T_x, X)]\}$ will be indirectly and non-selectively ordered in X! In the context of our synthesis, exactly the same kind of thing could happen in success and failures of classificatory separability.

Next, we entertain a general class of models within which both selective influence and separability can be theoretically lodged.

Accrual Halting Models

We have been making strides in recent years in building bridges between our strategies which have emphasized invariances associated with separability and independence and those constructed to identify mental architecture, workload capacity, and decisional stopping rules. The first, of course, belongs to the material above on classificatory behavior expressed in response probabilities whereas the latter is primarily associated with *RT*s, although there are subdivisions where this coupling does not hold.

Although the global mission is far from accomplished, we can point to several recent advances. Accordingly, we have generalized our nonparametric measure of workload capacity, $C(t)$, $t =$ time, which gauges performance with more than one process in operation against a prediction from one-process trials, based on an independent, unlimited capacity, parallel model. The new measure, $A(t)$, and called the assessment function, also rests on predictions from an independent, unlimited capacity, parallel model, but one which now allows errors to occur and, in fact, evaluates accuracy as well as response time in producing that measure (Townsend & Altieri, 2012).

More germane to our present situation, we have extended our general recognition theory of classification from response probabilities only to response probabilities and response times. With response times comes the question of architecture, so a particular class of architectures must be assumed in further developments. Our first theoretical efforts were founded on parallel processes. The individual processes were termed *accrual halting models*. All these models assume that activation (e.g., information) is being accrued in the relevant channels with a decision occurring when the first one reaches its own threshold. The definitions and theorems encompass all of the constituent models, which include most of the extant stochastic models of perception, action and decision making (see Townsend, Houpt, & Silbert, 2012 for details and more related studies and theories). The reader can find related developments along a somewhat different path, in Griffiths and colleagues' stochastic differential equations approach to a subclass of accrual models (Griffiths, Blunden, & Little, 2017).

This class of models seems to present an ideal environment for the bringing together of selective influence and classificatory separability and related notions. A sticking point would seem to be that up to now, almost all theorems and predictions relating to selective influence have been dedicated to response times and not accuracy (but see Schweickert et al., 2012 for exceptions). We are currently in the midst of theoretical developments intended to expand tests of architecture and stopping rule, founded on selective influence, to include both response probabilities as well as *RT*s (see dissertation of Yang, 2016).

How would these emerging tools be employed within parallel or serially arranged accrual halting models? If two physical values on two dimensions order distributions conditioned on, say, being correct, then tests of architecture and stopping rule could be straightaway applied. Needless to say, the assays of various types of independence and separability would also be available for implementation with the same set of data (Townsend et al., 2012).

In theory, other experimental factors, for instance, related to speed of processing of the dimensions, could be varied with selective influence in mind, besides the traditional values on the dimensions themselves.

And naturally, the general scope of such principles as classificatory separability and selective influence are relevant to any set of processes within a temporary or permanent processing system.

References

Algom, D., Eidels, A., Hawkins, R. X. D., Jefferson, B., & Townsend, J. T. (2015). *Features of response times: Identification of cognitive mechanisms through mathematical modeling*. *The Oxford handbook of computational and mathematical psychology*. New York, NY: Oxford University Press (pp. 63–98). Retrieved from http://ezproxy.lib.indiana.edu/login?. http://search.proquest.com/docview/1680462905?accountid=11620.

Ashby, F. G. (2014). Is state-trace analysis an appropriate tool for assessing the number of cognitive systems? *Psychonomic Bulletin & Review*, *21*(4), 935–946. http://dx.doi.org/10.3758/s13423-013-0578-x.

Ashby, F. G., & Soto, F. A. (2015). *Multidimensional signal detection theory*. *The Oxford handbook of computational and mathematical psychology*, *Vol. 13*.

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, *21*(2), 93–123. http://dx.doi.org/10.1016/0022-2496(80)90001-2.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179. http://dx.doi.org/10.1037/0033-295X.93.2.154.

Devaney, R. (1986). *An introduction to chaotic dynamical systems*. New York: Addison Wesley.

Dzhafarov, E. N. (2003). Selective influence through conditional independence. *Psychometrika*, *68*(1), 7–25. http://dx.doi.org/10.1007/BF02296650.

Dzhafarov, E. N., & Kujala, J. V. (2010). The joint distribution criterion and the distance tests for selective probabilistic causality. *Frontiers in Psychology: Quantitative Psychology and Measurement*, *1*, 151. http://dx.doi.org/10.3389/fpsyg.2010.00151.

Dzhafarov, E. N., & Kujala, J. V. (2012). Selectivity in probabilistic causality: Where psychology runs into quantum physics. *Journal of Mathematical Psychology*, *56*(1), 54–63. http://dx.doi.org/10.1016/j.jmp.2011.12.003.

Garner, W. R., & Morton, J. (1969). Perceptual independence: Definitions, models, and experimental paradigms. *Psychological Bulletin*, *72*(4), 233–259. http://dx.doi.org/10.1037/h0028024.

Griffiths, D. W., Blunden, A. G., & Little, D. R. (2017). Logical-rule based models of categorization: Using systems factorial technology to understand feature and dimensional processing. In D. R. Little, N. Altieri, M. Fifić, & C.-T. Yang (Eds.), *Systems factorial technology: A theory driven methodology for the identification of perceptual and cognitive mechanisms*. Elsevier (in this book).

Kadlec, H., & Townsend, J. T. (1992). Implications of marginal and conditional detection parameters for the separabilities and independence of perceptual dimensions. *Journal of Mathematical Psychology*, *36*(3), 325–374. http://dx.doi.org/10.1016/0022-2496(92)90027-5.

Kujala, J. V., & Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, *52*(2), 128–144. http://dx.doi.org/10.1016/j.jmp.2008.01.008.

Kujala, J. V., & Dzhafarov, E. N. (2010). Erratum to "Testing for selectivity in the dependence of random variables on external factors". *Journal of Mathematical Psychology*, *54*(4), 400. http://dx.doi.org/10.1016/j.jmp.2010.05.001.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147–180). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Chapter xiii, 523 pages. Retrieved from http://ezproxy.lib.indiana.edu/login?. http://search.proquest.com/docview/618227072?accountid=11620.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, *18*(2), 105–139. http://dx.doi.org/10.1016/0022-2496(78)90059-7.

Schweickert, R. (1982). The bias of an estimate of coupled slack in stochastic PERT networks. *Journal of Mathematical Psychology*, *26*, 1–12.

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

Silbert, N. H., & Thomas, R. D. (2013). Decisional separability, model identification, and statistical inference in the general recognition theory framework. *Psychonomic Bulletin & Review*, *20*(1), 1–20. http://dx.doi.org/10.3758/s13423-012-0329-4.

Soto, F. A., Vucovich, L., Musgrave, R., & Ashby, F. G. (2015). General recognition theory with individual differences: A new method for examining perceptual and decisional interactions with an application to face perception. *Psychonomic Bulletin & Review*, *22*(1), 88–111.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315. http://dx.doi.org/10.1016/0001-6918(69)90055-9.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, *28*(4), 363–400. http://dx.doi.org/10.1016/0022-2496(84)90007-5.

Townsend, J. T. (1990a). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, *108*(3), 551–567. http://dx.doi.org/10.1037/0033-2909.108.3.551.

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

Townsend, J. T. (1992). Chaos theory: A brief tutorial and discussion. In *From learning processes to cognitive processes: Vol. 2. Essays in honor of William K. Estes, Vol. 1: From learning theory to connectionist theory* (pp. 65–96). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Retrieved from http://ezproxy.lib.indiana.edu/login?. http://search.proquest.com/docview/618222203?accountid=11620.

Townsend, J. T., & Altieri, N. (2012). An accuracy-response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review*, *199*(3), 500–516.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan, & F. Restle (Eds.), In *Cognitive theory: Vol. III* (pp. 200–239). Hillsdale, NJ: Erlbaum Associates. Retrieved from http://www.indiana.edu/~psymodel/papers/towash78.pdf.

Townsend, J. T., & Ashby, F. G. (1983). A general discussion of equivalent and nonequivalent properties of serial and parallel systems and their models. In *The stochastic modeling of elementary psychological processes* (p. 481). Cambridge: Cambridge University Press. Chapter 15. Retrieved from http://www.indiana.edu/~psymodel/papers/Townsend_and_Ashby_Part3.pdf.

Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology*, *56*(6), 476–494. http://dx.doi.org/10.1016/j.jmp.2012.09.001.

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

Townsend, J. T., & Schweickert, R. (1985). Interactive effects of factors prolonging processes in latent mental networks. In G. d'Ydewalle (Ed.), In *XXIII international congress of psychology: Vol. 3. Cognition, information processing, and motivation* (pp. 255–276). Amsterdam: North Holland.

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, *33*(3), 309–327. http://dx.doi.org/10.1016/0022-2496(89)90012-6.

Townsend, J. T., & Spencer-Smith, J. B. (2004). Two kinds of global perceptual separability and curvature. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition*. Mahwah, NJ: Erlbaum.

Townsend, J. T., & Thomas, R. D. (1993). On the need for a general quantitative theory of pattern similarity. In *Foundations of perceptual theory* (pp. 297–368). Amsterdam: North-Holland/Elsevier Science Publishers.

Townsend, J. T., & Thomas, R. (1994). Stochastic dependencies in parallel and serial models: Effects on systems factorial interactions. *Journal of Mathematical Psychology*, *38*, 1–34.

Townsend, J. T., & Wenger, M. J. (2004). The serial-parallel dilemma: A case study in a linkage of theory and method. *Psychonomic Bulletin & Review*, *11*(3), 391–418. http://dx.doi.org/10.3758/BF03196588.

Wenger, M. J., & Ingvalson, E. M. (2003). Preserving informational separability and violating decisional separability in facial perception and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1106.

Wenger, M. J., & Townsend, J. T. (2001). Faces as gestalt stimuli: Process characteristics. In M. J. Wenger, & J. T. Townsend (Eds.), *Face perception and cognition*. Erlbaum Press.

Yang, H. (2016). *Survivor interaction contrast extended to include response accuracy* (Unpublished doctoral dissertation). Bloomington, IN: Indiana University.

Endnotes

1. The ways in which theory manifests change vs. invariance takes on highly recondite and powerful forms in 20th century physics. The prime example must be Emmy Noether's stupendous theorem which, with simplified language, demonstrated that for every symmetry in nature such as invariance under rotation, some type of physical quantities will be preserved in action over a path (see "Invariante Variations Probleme" [Invariant Variation Problems], *Nachr. D. König. Gesellsch. D. Wiss* (in German) (Göttingen: Math-phys. Klasse) 1918: 235–257. English translation by M.A. Tavel (1918); the interested reader will also delight in "Emmy Noether's Wonderful Theorem", charming and immensely informative, D.E. Neuenschwander, 2011).
2. There are curious features related to dependencies that can arise in nonlinear dynamics (such as chaos theory, Townsend, 1992; Devaney, 1986) but these lie outside our present concerns.
3. The question of probabilistic dependence arises in many guises, even in our own work, which must be sidestepped here. Starting references on those within our approaches are (a) Townsend and Ashby (1983) and (b) Ashby and Townsend (1986).
4. For instance, the main foundations of perceptual separability lie in General Recognition Theory (Ashby & Townsend, 1986) which itself has been cited in hundreds of papers (Ashby, 2014, personal communication). It undoubtedly figures in hundreds more in diverse studies of independence of perceptual dimensions (e.g., Garner & Morton, 1969). The notion of selective influence, lying as it does at the feet of the *additive factors method* (e.g., Sternberg, 1969) and its generalization, *systems factorial technology* (Townsend & Nozawa, 1995), has also been highly popular in experimental and methodological literature.
5. The concept of selective influence as stated here can be generalized to include more factors affect more processes and it is possible to differentiate between, say, random vectors, random variables and most generally, random entities (see, e.g., Dzhafarov & Kujala, 2010) but this simpler version will suffice for our purposes. And, since 1978 (Journal of Mathematical Psychology), Schweickert and colleagues have been investigating architectures which are more complex than the canonical serial vs. parallel arrangements (Schweickert, 1978; Townsend & Schweickert, 1985; see comprehensive review in Schweickert, Fisher, & Sung, 2012).

6. Similar statements apply to the notion of separability.
7. This curious recurrence of the same notion in palpably distinct settings, suggests that it may refer to a reasonably profound concept. That it has relevance to rather weighty conceptions is bolstered by Dzhafarov and colleagues' discovery of the relationship of marginal selectivity to logic associated with long-standing issues in quantum mechanics, including the famous Einstein–Podalsky–Rosen experiment and the linked Bell's inequality. However, this avenue lies outside our present concerns.