Reprinted from JOURNAL OF MATHEMATICAL PSYCHOLOGY
All Rights Reserved by Academic Press, New York and London

Vol. 38, No. 1, March 1994
Printed in Belgium

Stochastic Dependencies in Parallel and Serial Models: Effects on Systems Factorial Interactions

JAMES T. TOWNSEND AND ROBIN D. THOMAS

Indiana University

This paper examines the behavior of stochastically dependent serial and parallel processing models in the setting of a 2×2 factorial experiment. Interactions found in factorial experiments can provide insight into the underlying mental architecture operating in a given psychological task. Recent theoretical results classify mental networks according to the types of factorial interaction they predict when selectivity of the factors is assumed. When one allows this selectivity to break down through a stochastic dependence between processes, the characteristic patterns associated with distinct architectures are disturbed. We investigate the relationships among: (a) parallel and serial architectures, (b) positive, negative, and zero dependencies, and (c) types of mean reaction time factorial interaction. In particular, we show that in some cases, observable symptoms of a stochastic dependence arise. One of these, which we term a *single factor reversal* of mean processing times, arises as a result of a negative dependence under certain circumstances. Another characteristic of dependent systems that may contribute to their identifiability is that the interactions can be subadditive for some levels of the factors and superadditive at other levels. This change in contrast is not possible for independent serial and parallel models with selective influence.

Reaction time analysis has been a popular and important tool in the investigation of mental processing since the time of F. C. Donders' (1869) classic paper on this issue (see e.g., Ashby & Townsend, 1980; Luce, 1986). Donders outlined a method, today known as the *method of subtraction*, that is still practiced from time to time (e.g., Sternberg, 1966; Theios, 1973; Ashby, 1982; Ashby & Townsend, 1980; Vorberg, 1981; Gottsdanker & Shragg, 1985; Posner, 1978). Donders assumed that the mental processing occurring between a stimulus and a response is accomplished by *serially* arranged stages, or processes,[1] each of which could be inserted or removed as a result of experimental manipulations. He then designed a psychological

Reprint requests should be addressed to James T. Townsend, Dept. of Psychology, Indiana University, Bloomington, IN 47405.

[1] The meaning of the term *stage* when speaking of a serial system is obvious. However, when speaking of a parallel system the term becomes confusing. In much of Townsend's work (e.g., Townsend and Ashby, 1983) the term *stage* is used to delineate the duration between the successive completions of two elements undergoing processing. Because we are interested in the mechanisms doing the processing rather than what is being processed, we will use the terms *process* and *subsystem* to describe a distinct processing entity.

0022-2496/94 $6.00

task exploiting this assumption of *pure insertion*. By comparing the mean reaction time (RT) of one condition thought to include a particular process with the mean RT of another that did not include the process, he derived an estimate of the time consumed by the removed process. The strong assumptions of seriality and pure insertion needed for this method led to criticisms that limited its generality.

One hundred years later, Sternberg (1969) relaxed the assumption of pure insertion and proposed a method of studying mental architectures using a factorial design termed the *additive factor method*. Recall that a (complete) factorial design is an experimental design in which every level of every independent variable, or factor, is paired with every level of every other factor. An important feature of factorial designs is that they allow us to look at the interaction of variables. In the additive factor method, the desired finding would be that of non-interacting (i.e., additive) effects of experimental factors on mean RT. The method assumed, like Donders' strategy, that there are successive functional processes between a stimulus and a response whose durations are additive components of RT. Unlike Donders, Sternberg did not require the complete removal or insertion of a process for this method to be used. Rather, the method assumed that one could find experimental factors that *selectively* affect these serial processes by simply changing their respective durations. If the processing times were *stochastically independent* then, such factors would have an *additive* effect on RT in a factorial experiment.[2] That is, the effect of each factor on RT would be the same regardless of the levels of the other factors.

On the other hand, if a statistical interaction of the factors resulted, either the appropriate factors had not been found or the factors were affecting at least one stage in common. An interaction occurs if an effect of a given factor depends on the levels of the other factors. Factorial interactions can be of two types. In one case, combined effects of the factors can be less than the sum of their individual effects. This is termed *subadditivity*. If there are only two factors, subadditivity occurs when a mean RT decrease due to one factor is smaller, the larger the mean RT is for the other factor. The other type of interaction, *superadditivity*, can be defined analogously, but in the opposite direction. Figure 1 displays the three types of interactions (including additivity) in a typical factor graph. The assumption of seriality itself is not tested with the original additive factor method.

Investigations by Townsend (e.g., Townsend, 1971, 1972, 1976a, b) and Vorberg (1977) had earlier indicated broad classes of serial and parallel models capable of minicking each other's behavior, in many cases being mathematically equivalent. Even in the popular Sternberg memory scanning paradigm (Sternberg, 1966), widely believed to support serial processing, there were shown to exist mathematically identical and psychologically intuitive parallel models (Townsend, 1971, 1972).

[2] Mean RT additivity is possible without stochastic independence but it is not implied (Townsend and Ashby, 1983, Chap. 12).

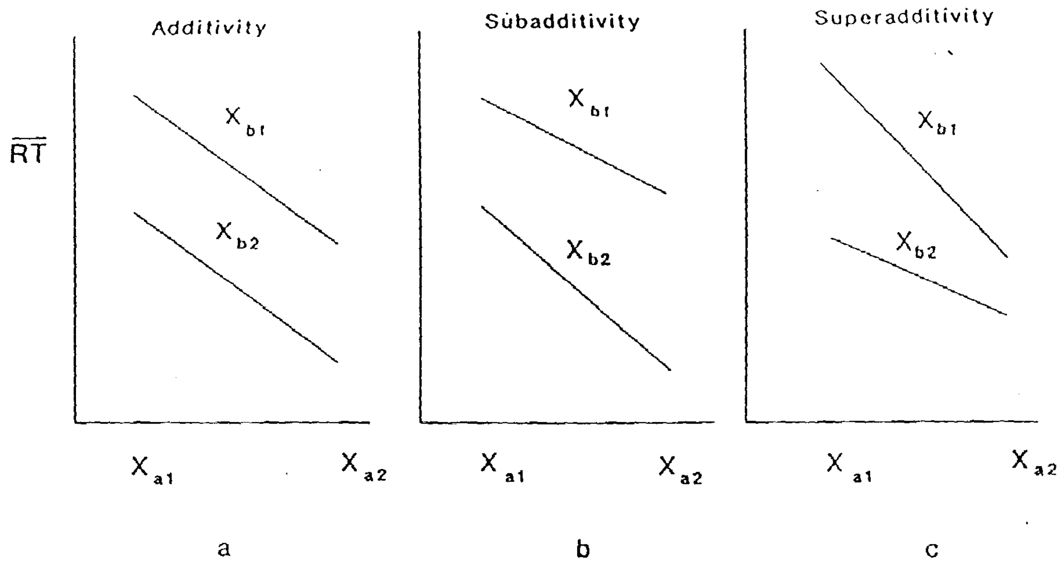

FIG. 1. A factor graph showing the three types of interaction: (a) additivity, (b) subadditivity, and (c) superadditivity.

Thus, it came as something of a pleasant surprise when generalizations of the additive factor method evidenced considerable promise of discriminating serial from parallel process models, as well as of diagnosing more general networks. Key early papers in the evolution of what we might now refer to as "systems factorial technology" are due to Schweickert (1977, 1978, 1980, 1982, 1983a, b, 1985) and Townsend and colleagues (Townsend and Piotrowski, 1981; Townsend and Ashby, 1983; Townsend, 1984; Townsend and Schweickert, 1985). Thus, the factorial methods joined an expanding set of tools with potential for distinguishing parallel from serial processes (reviewed in Townsend, 1990a) as well as for diagnosing more complex architectures. In particular, parallel processing is associated with subadditive interactions (e.g., Schweickert, 1978; Townsend and Ashby, 1983) as opposed to the additivity of typical serial systems. This study focuses entirely on parallel and serial structures. Other work in stochastic networks includes that of Fisher and Goldstein (1983), Fisher *et al.* (1985), and Goldstein and Fisher (1991).

As observed above, a highly central concept in factorial methodology is that of selective influence of each experimental factor on a specific psychological subprocess. The assumed existence of selective influence makes possible the derivation of theorems which associate distinct varieties of reaction time results with distinct classes of mental architectures. Little information has been available as to what happens when factorial selective influence is absent. Just how bad do things become? Is there any hope for systems identification under abrogation of selective influence? More specifically, can the failure of selective influence be observed? Is there any connection with the way in which selective influence fails and the type of factorial interaction?

In Theorem I, we first isolate an observable flag or symptom that is associated with a particular mode of breakdown of selective influence (the latter hereafter called "nonselective influence"). This flag may appear with either serial or parallel processing. Theorem II demonstrates that contrast sign (i.e., positive vs negative vs zero mean reaction time interaction) is logically independent of the *type* of stochastic dependence (i.e., positive vs negative). This absence of a relationship between contrast sign and dependence type permits the appearance of a second diagnostic dependence (and therefore indirect nonselective influence), as shown in Theorem III. The first flag of dependence associated with Theorem I, is revealed in Theorem IV to bear no relationship to contrast sign either. The final discussion places these findings within the general context of current experimentation and methodology and deliberates the potential for identifying indirect nonselective influence. We now further develop these concepts in detail.

A failure of selective influence could, of course, occur due to direct sources: Factor X_a, say, affects subsystem S_b in addition to S_a, the subsystem it is supposed to affect (see next section for details on notation). This is known as *direct* nonselective influence. However, a more subtle and interesting kind of failure of selective influence can occur through stochastic dependencies. This phenomenon is known as *indirect* nonselective influence (Townsend, 1984).

A strong dependence in a simple serial network could exist if, for example, the quality of output of a process is directly proportional to its processing duration and if the duration of a subsequent process tends to decrease with better input. Here, there would be a *negative* dependence between the two processes. When the first one takes longer to be completed, the second process has better input, thus requiring, on the average, less time.

Interference of facilitation effects that are manifested at a marginal, averaged level, could, but need not, be associated with intratrial dependencies. For instance, consider the well known Stroop effect where reading the name of a color printed in an incompatible ink hue causes interference relative to a neutral or compatible hue. It could be that on trials where the word is perceived very rapidly, greater interference is caused on the perception of the color than on trials when the word is perceived more slowly. The greater interference would presumably cause longer response times and thus generate a negative dependence between the color perception and reading times. See Logan and Zbrodoff (1979), MacLeod (1991), and Schweickert (1978, 1983a) for other accounts and more general discussion of Stroop effect phenomena.

Basic Terminology and Notation

This section formalizes many of the terms that were used above consistent with the usage adopted by Townsend and Ashby (1983), and Townsend (1984). The overall network in question may be called a *system*, if describing a real biological or psychological entity, or a *model*, if describing a mathematical structure used to make predictions about the system. Component processes in the network may be called *subprocesses* or *subsystems* of the system. Subprocesses may be arranged *serially*, that is, one immediately following another with no overlap in processing, or in *parallel*, that is, beginning simultaneously and ending randomly. It will be assumed in the present treatment that the serial systems are single path or, equivalently, fixed processing order. That is, it will be assumed that S_a always performs its job first, then S_b. This may be reasonable in some cognitive situations and less so in others. In any case, none of the central arguments that follow depend on this postulate, but it simplifies the notation somewhat. The present developments apply to experimental circumstances in which exhaustive processing is required. That is, both subprocesses must always complete their subtask before a response is made. The following notation is adopted:

S = total (real) system under consideration

S_i = a specific subsystem of S, $i = a, b, \ldots$

T = random variable for the processing time of the entire system, S, from $t = 0$ to completion of processing

T_i = random variable for the processing time of a particular subsystem, S_i, $i = a, b, \ldots$

X_i = experimentally manipulatable factor, considered as a set, often a non-negative numerical scale, associated with subsystem S_i, $i = a, b, \ldots$

f_i, F_i = probability density function and distribution on T_i, respectively, of subsystems arranged serially, $t_i \in [0, \infty)$, $i = a, b, \ldots$

g_i, G_i = density and distribution function on T_i, respectively, of subsystems arranged in parallel

$E_a(*)$, $E_b(*)$, $E_{ab}(*)$ = expectation taken relative to T_a, T_b, or both, respectively.

The probability functions are assumed to be functions of the experimental factors. That is, $f_a(t_a; x_a, x_b)$ would represent the density on T_a of a serial subsystem S_a, operating under the influence of $x_a \in X_a$, $x_b \in X_b$.

Using the above notation, we can now define precisely the notion of a parallel and an serial system as they are used in this paper. Because all of the results are for exhaustive processing, that assumption is implicit and need not be further repeated. The following definitions of parallel and serial systems do not yet impose selective influence. That will ensue below.

DEFINITION 1. A *parallel system* consists of subprocesses S_a and S_b with processing time random variables T_a and T_b, respectively, whose joint density is denoted by $g_{ab}(t_a, t_b; x_a, x_b)$, where $0 < t_a < \infty$ and $0 < t_b < \infty$. The joint distribution function is

$$P_{ab}(T_a \leqslant t_a, T_b \leqslant t_b; x_a, x_b) = G_{ab}(t_a, t_b; x_a, x_b) = \int_0^{t_a} \int_0^{t_b} g_{ab}(t'_a, t'_b; x_a, x_b)\, dt'_b\, dt'_a.$$

The total processing time T for the entire system is the longest processing time associated with the two processes, that is, $T = \max(T_a, T_b)$. The distribution on T is

$$P_{ab}(T \leqslant t; x_a, x_b) = P_{ab}(\max(T_a, T_b) \leqslant t; x_a, x_b) = G_{ab}(t_a, t_b; x_a, x_b)|_{\substack{t_a = t\\ t_b = t}}$$

One can find the expectation of T from

$$E_{ab}(T; x_a, x_b) = \int_0^{\infty} (1 - G_{ab}(t_a, t_b; x_a, x_b)|_{\substack{t_a = t\\ t_b = t}})\, dt$$

DEFINITION 2. A *serial system* consists of subsystems S_a and S_b such that the processing order is $\langle S_a, S_b \rangle$ with processing time random variables T_a and T_b, respectively, whose (conditional) densities are $f_a(t_a; x_a, x_b \mid t_b)$ and $f_b(t_b; x_a, x_b \mid t_a)$. The total processing time for the system is $T = T_a + T_b$.

The expectation of T can be found from

$$\begin{aligned} &E_{ab}(T_a + T_b; x_a, x_b) \\ &\quad = \int_0^{\infty} \int_0^{\infty} f_a(t_a; x_a, x_b)\, f_b(t_b; x_a, x_b \mid t_a)(t_a + t_b)\, dt_b\, dt_a \\ &\quad = E_a(T_a; x_a, x_b) + \int_0^{\infty} \int_0^{\infty} f_a(t_a; x_a, x_b)\, f_b(t_b; x_a, x_b \mid t_a)\, t_b\, dt_b\, dt_a \\ &\quad = E_a(T_a; x_a, x_b) + E_a[E_b(T_b; x_a, x_b \mid T_a)]. \end{aligned}$$

Experimental factors can exert three types of influence on a subsystem. While the following definitions are expressed in terms of the serial density functions, they hold for the parallel density functions as well. *Selective influence* exists if the factor directly affects only its own associated subsystem. In this case, the density for S_a would be a function only of x_a and similarly for S_b and x_b. We say factor X_b selectively affects subsystem S_b if the marginal density for T_b, the processing time of S_b, is a constant function of x_a for all values of t_b. In this case, we write $f_b(t_b; x_a, x_b) \equiv f_b(t_b; x_b)$.

DEFINITION 3. (Absolute) selective influence of factor X_a occurs if and only if $f_b(t_b; x_a, x_b) \equiv f_b(t_b; x_b)$. That is, the factor X_a does not influence the density of T_b.

Similarly, selective influence of factor X_b occurs if and only if $f_a(t_a; x_a, x_b) = f_a(t_a; x_a)$.

Direct nonselective influence describes the case when both factors directly affect at least one stage in common.

DEFINITION 4. Only direct nonselective influence of factor X_a occurs if and only if $f_b(t_b; x_a, x_b) \neq f_t(t_b; x_b)$, yet $f_b(t_b; x_a, x_b \mid t_a) \equiv f_b(t_b; x_a, x_b)$. Only direct nonselective influence of factor X_b occurs if and only if $f_a(t_a; x_a, x_b) \neq f_a(t_a; x_a)$, yet $f_a(t_a; x_a, x_b \mid t_b) \equiv f_a(t_a; x_a, x_b)$.[3]

That is, there is no influence via a stochastic time dependence of T_b on T_a, but X_a directly affects the processing time distribution on T_b. The third type of influence, *indirect nonselective influence*, is brought about by the existence of a stochastic dependence.

DEFINITION 5. Indirect nonselective influence of the factor X_a occurs if and only if $f_b(t_b; x_b \mid t_a) \neq f_b(t_b; x_b)$, $f_b(t_b; x_a, x_b \mid t_a) \equiv f_b(t_b; x_b \mid t_a)$, and $f_b(t_b; x_a, x_b) \neq f_b(t_b; x_b)$. Indirect nonselective influence of factor X_b occurs if and only if $f_a(t_a; x_a \mid t_b) \neq f_a(t_a; x_a)$, $f_a(t_a; x_a, x_b \mid t_b) \equiv f_a(t_a; x_a \mid t_b)$, and $f_a(t_a; x_a, x_b) \neq f_a(t_a; x_a)$.

The expressions in Definition 5 preclude direct, and implicate indirect, contamination of the "inappropriate" subprocess. Note that the fact that, on the marginals, $f_b(t_b; x_a, x_b) \neq f_b(t_b; x_b)$ implies that the stochastic contamination is actually revealed at the marginal density level. If marginal selectivity[4] does not hold and such a dependence exists, an indirect nonselective influence will occur on the mean processing times. It is logically possible to have both types of nonselective influence present in a given system.

We assume the condition encompassed in the following definition.

DEFINITION 6. Assume the absence of direct nonselective influence. Assume the potential of stochastic dependence in the sense that T_b may depend on T_a. Then, a *causal serial system* satisfies the condition that

$$f_{ab}(t_a, t_b; x_a, x_b) = f_a(t_a; x_a)\, f_b(t_b; x_b \mid t_a).$$

[3] In this definition, we have explicitly prohibited any stochastic dependence between the processing times T_a and T_b. We could have, however, allowed a dependence to exist as long as this dependence was not detectable at the marginal density level. This more general condition is called marginal selectivity (e.g., Townsend and Schweickert, 1989) and is defined precisely in footnote 4.

[4] *Marginal selectivity* of factor X_a holds if $\int_0^\infty f_b(t_b; x_b \mid t_a) f_a(t_a; x_a)\, dt_a = f_b(t_b; x_b)$. Note, in general, this definition allows $f_b(t_b; x_b \mid t_a) \neq f_b(t_b; x_b)$, but the influence of X_a through the stochastic dependence of the random variable T_b on T_a is washed out after integrating over values of t_a. Marginal selectivity of factor X_b can be defined analogously.

Note that this does not imply $f_b(t_b; x_b) f_a(t_a; x_a \mid t_b) = f_{ab}(t_a, t_b; x_a, x_b)$. Nor does it imply $\int_0^\infty f_a(t_a; x_a) f_b(t_b; x_b \mid t_a)\, dt_a = f_b(t_b; x_b)$. However, Definition 6 does imply $\int_0^\infty f_a(t_a; x_a) f_b(t_b; x_b \mid t_a)\, dt_b = f_a(t_a; x_a)$. Thus, imposition of serial causality implies selective influence for the first subsystem, S_a, but allows the possibility of indirect nonselective influences on S_b. Although the theorems below could be proven without this assumption, we think it is important to represent our model classes in as realistic a fashion as currently possible. In sum, either type of contamination can occur on S_b but only direct influences could impinge on S_a (but are excluded from consideration here). It is indirect nonselective influence that is the topic of this work. As an example, consider the following conditional densities for a serial system with subsystems S_a and S_b, taken from Townsend (1984):

$$S_a : f_a(t_a; x_a) = x_a e^{-x_a t_a}, \qquad t_a > 0$$

$$S_b : f_b(t_b; x_b \mid t_a) = (x_b \mid t_a)\, e^{-(x_b t_b)/t_a}, \qquad t_a > 0,\ t_b > 0.$$

Note that there are no direct nonselective influences in that each factor, x_i, appears only in its respective exponential density function, f_i, $i = a, b$. An indirect nonselective influence of factor x_a is present in the density function of S_b since the rate parameter contains the first processing time, t_a, and $f_a(t_a)$ is a function of x_a.

Now we need to define precisely what is meant by a factorial interaction. Let T_{ij} be the total processing time when factor X_a is at level x_{ai} and factor X_b is at level x_{bj}. A statistic used in 2×2 factorial experiment to evaluate the nature of an interaction is the difference of mean differences or *mean contrast* in terms of the sample means, $\bar{T}_{ij}$,

$$\Delta^2 \bar{T} = \bar{T}_{22} - \bar{T}_{21} - \bar{T}_{12} + \bar{T}_{11}.$$

This is analogous to the second-order theoretical difference of expectations

$$\Delta^2 E(T_{ij}) = E(T_{22}) - E(T_{21}) - [E(T_{12}) - E(T_{11})]. \qquad (1)$$

In terms of model predictions, we say that an interaction effect is subadditive (superadditive) for a particular choice of factor levels if this second-order difference, or contrast sign, is negative (positive) for that choice of factor levels. If the second-order difference is zero, there is no interaction; that is, the factors are additive. Sometimes, it may be convenient to use continuous derivatives, if the probability functions are differentiable, in a manner completely analogous to the above finite difference results. In these cases we are concerned with the second-order mixed partial derivative $(\partial^2/\partial x_a\, \partial x_b)\, E_{ab}(T)$ when determining the nature of an interaction. Another related expression that is used in this paper is the operator Δ_{xi}, for $i = a, b$.

This is defined as the first-order difference, usually of expected processing times, with respect to the particular factor x_i.[5]

Ignoring how a factor affects another subsystem, we can discuss how a factor affects its own subsystem. Throughout this paper, it is assumed that there are two experimental factors that each tend to speed up two subsystems. This type of factor effect can be manifested in several ways. The simplest, and weakest, way is to assume that the factors order the mean processing times of the subsystems. A slightly stronger assumption is that the factors order the distributions of the processing times for their respective subsysems. Formally, this assumption states that, when X_a goes from x_{a1} to x_{a2}, where $x_{a2} > x_{a1}$, $F_a(t_a; x_{a2}) \geqslant F_a(t_a; x_{a1})$.[6] Distribution ordering implies that the means are ordered as well. A yet stronger assumption that can be made concerning the factor effects is the condition that there exists a *single crossover* for the densities associated with the factor levels. This, in turn, implies that the distributions are ordered, but the converse is not true. A single crossover exists if, for $x_{a2} > x_{a1}$, there is some time t^*, such that $f_a(t_a; x_{a2}) > f_a(t_a; x_{a2})$ for all $t < t^*$, $f_a(t_a; x_{a2}) < f_a(t_a; x_{a1})$ for all $t > t^*$, and $f_a(t^*; x_{a1}) = f_a(t^*; x_{a2}) > 0$. These statements also hold for the probability functions associated with subsystem S_b. (Further, we require, on occasion, that the speed-up hold in the conditional distributions, as in $P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t) > P(T_a \leqslant t; x_{a1} \mid T_b \leqslant t)$ when $x_{a2} > x_{a1}$.) The above assertions were established in Townsend and Ashby (1978, 1983), work that led to a full set of dominance relations (Townsend, 1990b). The latter two assumptions are employed in the proofs that follow on an "as needed" basis.

Concerning the nature of the stochastic time dependencies, there are several approaches to take as well. For most of our purposes, if a dependence exists, we simply assume that the conditional expectation of one processing time is a strict monotonic function of the processing time of the other subsystem in some interval.

[5] For an arbitrary function, h, of two variables, u and v, and for fixed $t > 0$, and fixed v, the first-order difference with respect to the variable u is defined to be $\Delta_u h(u, v) = h(u + t, v) - h(u, v)$. The first-order difference with respect to the variable v is defined analogously. In 2×2 factorial experiments with the variables being fixed factors, the variables u and v are assumed to be binary valued so that $u = \{u_1, u_2\}$ and $v = \{v_1, v_2\}$. We are assuming these sets are ordered so that $u_2 > u_1$ and $v_2 > v_1$. Here, then the first-order difference becomes $\Delta_u h(u, v) = h(u_2, v) - h(u_1, v)$. Also, $\Delta_v h(u, v) = h(u, v_2) - h(u, v_1)$. Now, the second-order difference, Δ^2, is just the first-order difference with respect to the variable v of the first-order difference with respect to the variable u. That is,

$$\begin{aligned}\Delta^2 h(u, v) &= \Delta_v(\Delta_u h(u, v)) \\ &= \Delta_v(h(u_2, v) - h(u_1, v)) \\ &= (h(u_2, v_2) - h(u_1, v_2)) - (h(u_2, v_1) - h(u_1, v_1)) \\ &= h(u_2, v_2) - h(u_1, v_2) - h(u_2, v_1) + h(u_1, v_1),\end{aligned}$$

which is how the second-order difference was defined earlier. Notice that $\Delta_v(\Delta_u) = \Delta_u(\Delta_v)$ so that Δ^2 is unambiguous.

[6] Again, we are ignoring for the moment the factor X_b. We can do this since in the following definitions we focus on how factor X_a influences its own subsystem, S_a, for a *fixed* level of X_b. Thus, in the notation that follows, although X_b does not appear, selective influence is not a necessary requirement for these definitions and is, *in fact*, an orthogonal issue.

In the case of a positive dependence of T_b on T_a, for example, this assumption states that $E_b(T_b; x_b \mid T_a = t_{a1}) < E_b(T_b; x_b \mid T_a = t_{a2})$ for all $t_{a1} < t_{a2}$ in some nonnegative interval of T_a. For a negative dependence, we would have $E_b(T_b; x_b \mid T_a = t_{a1}) < E_b(T_b; x_b \mid T_a = t_{a2})$ whenever $t_{a1} > t_{a2}$. In some cases (e.g., the parallel case of Theorem I) we require that the conditional time cumulative distributions be ordered with respect to the other processing time. That is, $P(T_i \leqslant t; x_i \mid T_j = t_j)$ is a monotonic function of t_j $(i \neq j)$. As stated earlier, this distribution ordering implies that the conditional expectations are ordered as well. Now that we have the necessary tools in hand, let us proceed to the business of dependent model behavior.

The Single Factor Reversal Phenomenon

There are two types of factorial behaviors that are indicative of simple dependent models. The first, the *single factor reversal* phenomenon, is described in this section. In addition, serial and parallel dependent models can exhibit a *contrast change*; that is, they can alter their contrast sign as a function of the factor levels. In other words, they can go from behaving subadditively to behaving superadditively. Parallel and serial systems obeying marginal selective influence cannot do this (see Theorem II below). This feature is described in detail in the following section.

In a typical additive factor model, if the factor effects are defined as above (i.e., they both tend to speed up processing) and the processes are stochastically independent, a conventional ordering of the mean RTs results. For example, if our two factors, X_a and X_b, when increased tended to speed up processing of their respective subsystems, the following order of the mean processing times for the four conditions is induced if selective influence holds:

> $E(T_{11}) > E(T_{12})$, $E(T_{21}) > E(T_{22})$, where $E(T_{ij})$ signifies the overall expected processing time of the entire system, when factor X_a is at level x_{ai}, and factor X_b is at level x_{bj}, $i, j = 1, 2$.

The relationship of the middle two expectations to each other is not determined. A reversal in the ordering of the means with respect to a single factor, say X_a, occurs if $E(T_{21}) > E(T_{11})$ or $E(T_{22}) > E(T_{12})$, or both. It is shown in Theorem I that a necessary condition for a single factor reversal is that there exists a negative depedence, at least at the level of the conditional expectations, between the two subprocesses influenced by the experimental factors. One might argue that such a reversal may be due, instead, to a change in the manner in which an experimental factor affects processing. That is, a factor may act to speed up processing for some of its levels while it slows down such processing for others. *Throughout this paper, we assume that experimental factors always have monotonic effects on processing speed and that any nonselective influence is the result of a stochastic dependence and is not direct.* Otherwise, specific assumptions are outlined in the statement of the theorems. It is not unreasonable to postulate that increasing stimulus brightness, for example, always leads to enhanced performance of the subsystem responsible for its perception. Such assumptions concerning factor effects are ubiquitous in the literature and would presumably apply in the situations that are described here.

For the following theorem, it does not matter whether processing is parallel or serial; consequently, a proof is offered for both.

THEOREM I.a. *For causal processing, let increasing the level of each factor impose a single crossover in each of the (conditional) densities on the process durations,* T_a *and* T_b. *Also, direct nonselective influence does not occur. The presence of a single factor reversal implies that there exists a negative dependence of the time required by* S_b, T_b, *on the time taken by* S_a, T_a *is some subinterval of* $T_a \in [0, \infty)$. *Stochastic dependence is at the level of conditional expectations. That is, there exists an indirect nonselective influence of factor* X_a *on subsystem* S_b *as per the negative dependence.*

Proof. The proof is a modification of a line of argument used by Townsend and Schweickert (1985, Theorem 2). The presence of a single factor reversal with respect to the factor, say X_a, for a particular level of the other factor, X_b, requires that $\Delta_{xa} E_{ab}(T_a + T_b; x_a, x_b) = E_{ab}(T_a + T_b; x_{a2}, x_b) - E_{ab}(T_a + T_b; x_{a1}, x_b) > 0$ for some $x_{a2} > x_{a1}$ and fixed x_b. Recall that $E_{ab}(T_a + T_b; x_a, x_b) = E_a(T_a; x_a) + E_a[E_b(T_b; x_b \mid T_a)]$, so

$$\Delta_{x_a} E_{ab}(T_a + T_b; x_a, x_b) = \Delta_{x_a} E_a(T_a; x_a) + \Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)]. \qquad (2)$$

The first term on the right-hand side of (2) is always negative because we have assumed that the factors speed up processing of their respective subsystems. Therefore, if the overal expression is to be positive, the second term on the right is necessarily positive. In order to prove that this requires that a negative dependence is present, we show that any other type of dependence (i.e., positive dependence or independence) for all T_a leads to a contradiction.

As stated above, we assume that X_a exerts its influence on S_a by rendering a single crossover of its processing time densities $f_a(t_a; x_{a2})$ and $f_a(t_a; x_{a1})$ at t^* such that whenever $x_{a2} > x_{a1}$,

$$f_a(t_a; x_{a2}) - f_a(t_a; x_{a1}) > 0 \qquad \text{when} \quad 0 \leq t_a < t^*,$$

$$f_a(t_a; x_{a2}) - f_a(t_a; x_{a1}) < 0 \qquad \text{when} \quad t_a > t^*,$$

and

$$f_a(t_a; x_{a2}) = f_a(t_a; x_{a1}) > 0 \qquad \text{when} \quad t_a = t^*.$$

Now,

$$\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] = \int_0^{\infty} [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})]\, E_b(T_b; x_b \mid T_a = t_a)\, dt_a. \qquad (3)$$

First, assume that T_a and T_b are independent. Then, $E_b(T_b; x_b \mid T_a = t_a) = E_b(T_b; x_b)$ so that (3) becomes

$$\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] = \int_0^\infty [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})] \, E_b(T_b; x_b) \, dt_a$$
$$= E_b(T_b; x_b) \int_0^\infty [f_a(t_a; x_{a2})] \, dt_a$$
$$= 0.$$

But this contradicts our requirements that $\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] > 0$ for the single factor reversal to occur.

Next, assume that T_b depends positively on T_a over all T_a in the sense that $E_b(T_b; x_b \mid T_a)$ is a strictly monotonic increasing function of T_a. Again,

$$\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] = \int_0^\infty [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})] \, E_b(T_b; x_b \mid T_a = t_a) \, dt_a.$$

For simplicity, write

$$a_1(t') = \begin{cases} f_a(t'; x_{a2}) - f_a(t'; x_{a1}) & 0 \leqslant t' \leqslant t^* \\ 0 & t' > t^* \end{cases}$$
$$a_2(t') = \begin{cases} f_a(t'; x_{a1}) - f_a(t'; x_{a2}) & t' > t^* \\ 0 & 0 \leqslant t' \leqslant t^* \end{cases}$$
$$c(t') = E_b(T_b; x_b \mid T_a = t').$$

Now, substituting into (3),

$$\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] = \int_0^{t^*} a_1(t') \, c(t') \, dt' - \int_{t^*}^\infty a_2(t') \, c(t') \, dt'.$$

Recall that $c(t')$ is a monotonic increasing function of t'. Note also that $\int_0^{t^*} a_1(t') \, dt' = \int_{t^*}^\infty a_2(t') \, dt'$. By the mean value theorem and strict monotonicity of $c(t')$, there exists a δ sufficiently close to 0 that the above conditions imply

$$\int_0^{t^*} a_1(t') \, c(t') \, dt' < c(t^*) \int_0^{t^*} a_1(t') \, dt'$$
$$< c(t^* + \delta) \int_{t^*}^\infty a_2(t') \, dt'$$
$$< \int_{t^*}^\infty a_2(t') \, c(t') \, dt'$$
$$\rightarrow \int_0^{t^*} a_1(t') \, c(t') \, dt' < \int_{t^*}^\infty a_2(t') \, c(t') \, dt'$$
$$\rightarrow \int_0^{t^*} a_1(t') \, c(t') \, dt' - \int_{t^*}^\infty a_2(t') \, c(t') \, dt' < 0$$
$$\rightarrow \Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] < 0,$$

which contradicts the fact that $\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)]$ is positive due to the single factor reversal. One can trace through the above argument to confirm that negative dependence on some interval is consistent with the reversal. In that case $c(t')$ is a monotonic decreasing function of t' in that interval and the inequalities would be reversed as needed. Note that a sufficient condition for a single factor reversal is that $\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] > 0$ and that

$$\Delta_{x_a} E_a[E_b(T_b; x_b \mid T_a)] > |\Delta_{x_a} E_a(T_a; x_a)|.$$

THEOREM I.b. *For parallel processing, let increasing the level of each factor impose single crossovers of the (conditional) densities of the process durations, T_a and T_b. Also, direct nonselective influence does not occur. The presence of a single factor reversal implies that there exists a negative dependence on the two subsystems, S_a and S_b. The stochastic dependence is at the level of the conditional distribution function. That is, there exists an indirect nonselective influence of one factor on the other subsystem as per the negative dependence.*

Proof. For parallel models, the expected exhaustive processing time is $E_{ab}[\max(T_a, T_b); x_a, x_b] = \int_0^\infty [1 - P(T_a \leqslant t, T_b \leqslant t; x_a, x_b)]\, dt$. Again, from the definition of a single factor reversal with respect to, say X_a, we know that $\Delta_{x_a} E_{ab}[\max(T_a, T_b); x_a, x_b] > 0$. Now, for a fixed x_b,

$$\Delta_{x_a} E_{ab}[\max(T_a, T_b); x_a, x_b]$$
$$= E_{ab}[\max(T_a, T_b); x_{a2}, x_b] - E_{ab}[\max(T_a, T_b); x_{a1}, x_b] > 0.$$

So,

$$\int_0^\infty [1 - P(T_a \leqslant t, T_b \leqslant t; x_{a2}, x_b)]\, dt$$
$$- \int_0^\infty [1 - P(T_a \leqslant t, T_b \leqslant t; x_{a1}, x_b)]\, dt > 0,$$

which leads to

$$\int_0^\infty [P(T_a \leqslant t, T_b \leqslant t; x_{a2}, x_b) - P(T_a \leqslant t, T_b \leqslant t; x_{a1}; x_b)]\, dt < 0. \qquad (4)$$

Recall that because the only possible influences from S_b to S_a occur from stochastic dependencies, we can write $P(T_a \leqslant t, T_b \leqslant t; x_a, x_b)$ as

$$P(T_a \leqslant t; x_a \mid T_b \leqslant t)\, P(T_b \leqslant t; x_a, x_b),$$

where $P(T_b \leqslant t; x_a, x_b) = \int_0^\infty P(T_b \leqslant t; x_b \mid T_a = t_a) g(t_a; x_a)\, dt_a$. So now, (4) becomes

$$\int_0^\infty \{P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t)\, P(T_b \leqslant t; x_{a2}, x_b) - P(T_a \leqslant t; x_{a1} \mid T_b \leqslant t)\, P(T_b \leqslant t; x_{a1}, x_b)\}\, dt < 0. \tag{5}$$

Adding and subtracting $P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t)\, P(T_b \leqslant t; x_{a1}, x_b)$, we get

$$\int_0^\infty \{P(T_a \leqslant t; x_{a2}, \mid T_b \leqslant t)[P(T_b \leqslant t; x_{a2}, x_b) - P(T_b \leqslant t; x_{a1}, x_b)] + [P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t) - P(T_a \leqslant t; x_{a1} \mid T_b \leqslant t)]\, P(T_b \leqslant t; x_{a1}, x_b)\}\, dt < 0;$$

that is,

$$\int_0^\infty \{P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t)\, \Delta_{x_a}[P(T_b \leqslant t; x_a, x_b)] + \Delta_{x_a}[P(T_a \leqslant t; x_a \mid T_b \leqslant t)]\, P(T_b \leqslant t; x_{a1}, x_b)\}\, dt < 0. \tag{6}$$

Because the terms $P(T_a \leqslant t; x_{a2} \mid T_b \leqslant t)$ and $P(T_b \leqslant t; x_{a1}, x_b)$ are probabilities, they are always nonnegative. Also, $\Delta_{x_a}[P(T_a \leqslant t; x_a \mid T_b \leqslant t)]$ is always positive due to our assumption that the factors tend to speed up processing of their subsystems. Therefore, the only means by which the above integral (6) can be negative is to have $\Delta_{x_a} P(T_b \leqslant t; x_a, x_b) < 0$. All that remains to be shown is that this requires a negative dependence between T_b and T_a to exist somewhere. We again employ the method of contradiction.

Now,

$$\Delta_{x_a} P(T_b \leqslant t; x_a, x_b) = \int_0^\infty [g_a(t_a; x_{a2}) - g_a(t_a; x_{a1})]\, P(T_b \leqslant t; x_b \mid T_a = t_a)\, dt_a, \tag{7}$$

where $x_{a2} > x_{a1}$. We again require the assumption that X_a orders the distributions, G_a, by imposing a single crossover of the densities, g_a, at t^*.

In analogy to the serial derivation, define

$$a_1(t') = \begin{cases} g_a(t'; x_{a2}) - g_a(t'; x_{a1}) & 0 \leqslant t' \leqslant t^* \\ 0 & t' > t^* \end{cases}$$

$$a_2(t') = \begin{cases} g_a(t'; x_{a1}) - g_a(t'; x_{a2}) & t' > t^* \\ 0 & 0 \leqslant t' \leqslant t^* \end{cases}$$

and

$$c(t') = P(T_b \leqslant t; x_b \mid T_a = t').$$

Substituting into (7), $\Delta_{x_a} P(T_b \leqslant t; x_a, x_b)$ becomes

$$\int_0^{t^*} a_1(t')\, c(t')\, dt' - \int_{t^*}^{\infty} a_2(t')\, c(t')\, dt'.$$

Suppose that T_b and T_a were positively correlated everywhere in the sense that $P(T_b \leqslant t; x_b \mid T_a = t')$ is a monotonically decreasing function of t'. That is, the probability that T_b would be less than or equal to some time t given a large T_a would be less than that probability given a small T_a. Note, again that $\int_0^{t^*} a_1(t')\, dt' = \int_{t^*}^{\infty} a_2(t')\, dt'$. As in Theorem Ia, there exists a δ sufficiently close to 0 that the above conditions imply

$$\int_0^{t^*} a_1(t')\, c(t')\, dt' > c(t^*) \int_0^{t^*} a_1(t')\, dt'$$

$$> c(t^* + \delta) \int_{t^*}^{\infty} a_2(t')\, dt'$$

$$> \int_{t^*}^{\infty} a_2(t')\, c(t')\, dt'$$

$$\rightarrow \int_0^{t^*} a_1(t')\, c(t')\, dt' > \int_{t^*}^{\infty} a_2(t')\, c(t')\, dt'$$

$$\rightarrow \int_0^{t^*} a_1(t')\, c(t')\, dt' - \int_{t^*}^{\infty} a_2(t')\, c(t')\, dt' > 0$$

$$\rightarrow \Delta_{x_a} \int_0^{\infty} P(T_b \leqslant t; x_b \mid T_a = t_a)\, g_a(t_a; x_a)\, dt_a > 0$$

$$\rightarrow \Delta_{x_a} P(T_b \leqslant t; x_a, x_b) > 0.$$

But we started with $\Delta_{x_a} P(T_b \leqslant t; x_a, x_b) < 0$ from the single factor reversal. Thus, positive dependence also leads us to a contradiction.

It is easy to show that if we assume independence, then

$$\Delta_{x_a} P(T_b \leqslant t; x_a, x_b) = 0,$$

again leading us to a contradiction by a similar argument used in the serial proof. Therefore, a necessary condition for a single factor reversal to occur in a parallel model is that there exists a negative dependence between the subprocesses at the level of the conditional distribution functions. ∎

A sufficient condition for a reversal is that both

$$\Delta_{x_a} P(T_b \leqslant t; x_a, x_b) < 0$$

and

$$|\Delta_{x_a} P(T_b \leqslant t; x_a, x_b)| > \frac{[\Delta_{x_a} P(T_a \leqslant t; x_a \mid T_b \leqslant t)]\, P(T_b \leqslant t; x_a, x_b)}{P(T_a \leqslant t; x_a \mid T_b \leqslant t)},$$

for all $t > 0$, and for the denominator always > 0.

To illustrate the concept of the single factor reversal phenomenon we offer a simple serial model with a negative dependence between S_b and S_a. It is important to note that X_a does *not* directly influence subsystem S_b. Any contamination of S_b's overall processing time, T_b, comes about *only* through the stochastic dependence. In the parlance of Definitions 4 and 5, we have indirect nonselective influence but not direct nonselective influence. Consider a system whose probability densities on the processing times for each subsystem are as follows:

$$S_a : f_a(t_a; x_a) = x_a e^{-x_a t_a}$$

$$S_b : f_b(t_b; x_b \mid t_a) = e^{t_a} x_b e^{-e^{t_a} x_b t_b}.$$

Observe the dependence expressed in the rate of S_b. As t_a grows larger, processing of S_b speeds up, thereby creating a negative correlation between the two subsystems. The expected processing times for each subsystem are

$$S_a : E_a(T_a; x_a) = 1/x_a,$$

$$S_b : E_a[E_b(T_b; x_b \mid T_a)] = x_a/[x_b(x_a + 1)].$$

The total expected processing time for the system is just the sum of the two expectations

$$E_{ab}(T_a + T_b; x_a, x_b) = 1/x_a + x_a/[x_b(x_a + 1)].$$

A sufficient condition for X_a to reorder the means is that $\partial/\partial x_a[E_{ab}(T; x_a, x_b)]$ be positive for that choice of factor levels. This can happen if we let $x_{a1} = 1$, $x_{a2} = 2$, $x_{b1} = \frac{1}{8}$, and $x_{b2} = \frac{1}{4}$. The expectations for the four conditions are then

Condition	$E(T)$
(1, 1)	5
(1, 2)	3
(2, 1)	35/6
(2, 2)	19/6

With these factor levels the means are reordered such that $E(T_{21}) > E(T_{11}) > E(T_{22}) > E(T_{12})$. A 2×2 factorial graph that exhibits this reversal is shown in Fig. 2.

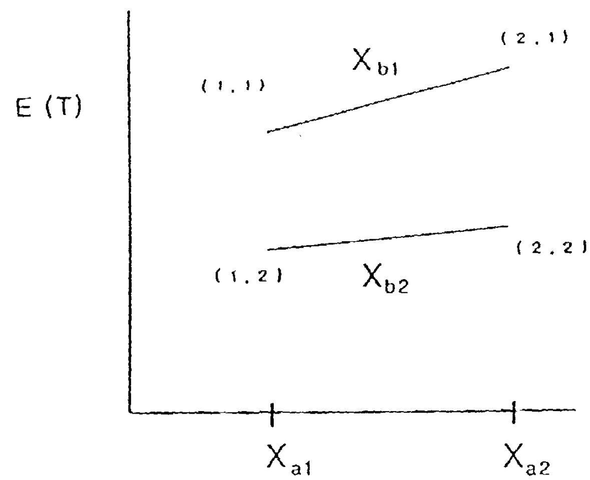

FIG. 2. A factor graph illustrating a single factor reversal, here with respect to the factor X_a.

It is not always the case that there is a reversal with this model. For instance, if $x_{a1}=1$, $x_{a2}=2$, $x_{b1}=1$, $x_{b2}=2$, then the expectations for the four conditions are

Condition	$E(T)$
(1, 1)	3/2
(1, 2)	5/4
(2, 1)	7/6
(2, 2)	5/6

Here, a typical ordering of the means occurs; that is, $E(T_{11}) > E(T_{12}) > E(T_{21}) > E(T_{22})$. This model is discussed again in later sections.

Logical Independence of Contrast and Type of Stochastic Dependence: Empirical Implications

A result that we initially found counterintuitive, illustrated in the first of the following theorems, is that the type of factorial interaction, or contrast sign, is logically independent of the type of dependence (positive or negative) inherent in the model. It had seemed to us that subadditivity might be associated with negative dependence and superadditivity with positive dependence. Again, note that direct nonselective influence (see Definition 4) is *not* allowed in the following theorems.

THEOREM II. *Assume that only indirect nonselective influence is in force. In both (causal) serial and parallel models, contrast sign is logically independent of the sign of the stochastic dependence producing the indirect nonselective influence.*

Proof. We may prove that these two concepts are logically independent by constructing the following models for both the serial and parallel case: (1) positively dependent and superadditive, (2) positively dependent and additive, (3) positively dependent and subadditive, (4) negatively dependent and subadditive, (5) negatively dependent and additive, and (6) negatively dependent and superadditive.

(*Serial*)

(1) *Positively dependent and superadditive.* A serial model satisfying the properties of this case has already been investigated in Townsend (1984, p. 371–372). Its density functions were given above to introduce indirect nonselective influence. For convenience, we restate the densities here, and refer the reader to the above paper for the proof of superadditivity. The density functions are

$$S_a : f_a(t_a; x_a) = x_a e^{-x_a t_a}, \qquad x_a > 0,\ t_a > 0$$

$$S_b : f_b(t_b; x_b \mid t_a) = (x_b/t_a)\, e^{-(x_b t_b)/t_a}, \qquad x_b > 0,\ t_a > 0,\ t_b > 0.$$

(2) *Positively dependent and additive.* Suppose that there exists a third subsystem, S_c, in addition to S_a and S_b, that plays the role of a capacity source for the two subsystems, consequently affecting their probability densities. Let the random variable C denote the amount of capacity available. Then, the three probability functions are

$$f_a(t_a; x_a \mid c) = c x_a e^{-c x_a t_a}, \qquad P_c(C) = \tfrac{1}{2} \quad \text{for} \quad C = 2$$

$$f_b(t_b; x_b \mid c) = c x_b e^{-c x_b t_b}, \qquad P_c(C) = \tfrac{1}{2} \quad \text{for} \quad C = 4.$$

Now,

$$\begin{aligned} E_{ab}(T_a + T_b; x_a, x_b) &= E_c[E_{ab}(T_a + T_b; x_a, x_b \mid C)] \\ &= E_c\left[\left(\frac{1}{C}\right)\left(\frac{1}{x_a} + \frac{1}{x_b}\right)\right] \\ &= \left(\frac{3}{8}\right)\left(\frac{1}{x_a} + \frac{1}{x_b}\right), \end{aligned}$$

for which the second mixed partial derivative $\partial^2/\partial x_a\, \partial x_b = 0$.

The fact that this model exhibits positive dependence can probably be seen in that both subsystems, S_a and S_b, will have a large capacity or a small capacity on any given trial. However, to be sure that this is true, one can compare the conditional probability that, say T_b is greater than some time t_b, given that T_a is greater than some time t_a, versus the marginal probability that T_b is greater than t_b. If there exists a positive dependence, at the level of the conditional probabilities, then the conditional probability should be larger than the marginal probability.

That is,

$$P(T_b > t_b \mid T_a > t_a) > P(T_b > t_b),$$

where

$$P(T_b > t_b \mid T_a > t_a) = \frac{P(T_b > t_b \cap T_a > t_a)}{P(T_a > t_a)}.$$

A proof along these lines for a similar model exists in Townsend and Ashby (1983, Chap. 12, pp. 368–370) and hence is not repeated here, but is available from the authors upon request.

(3) *Positively dependent and subadditive.* Let the densities for the subsystem S_a be defined as

$$S_a : f_a(t_a; x_{a1}) = \frac{1}{4} e^{-(1/4)t_a}$$

$$f_a(t_a; x_{a2}) = e^{-t_a}$$

$$S_b : f_b(t_b; x_b \mid t_a) = \frac{1}{R(t_a, x_{bi})} e^{(1/R(t_a, x_{bi}))t_b},$$

where $E_b(T_b; x_{bi} \mid T_a) = R(t_a, x_{bi})$ where $i = 1, 2$. Now we construct $R(t_a, x_{bi})$ so that there exists a positive dependence between T_b and T_a. Let the expected processing time for T_b for each level of X_b be given as

$$E_b(T_b; x_{b1} \mid T_a = t_a) = \begin{cases} \sqrt{2t_a} & 0 \leqslant t_a \leqslant 1.8 \\ t_a + 0.097 & t_a > 1.8, \end{cases}$$

and

$$E_b(T_b; x_{b2} \mid T_a = t_a) = \begin{cases} \frac{1}{2}t_a^2 & 0 \leqslant t_a \leqslant 1.8 \\ t_a - 0.18 & t_a > 1.8. \end{cases}$$

The positive dependence is reflected in the fact that for both levels of X_b, the expected processing time for S_b is an increasing function of T_a. Because the densities and the conditional expectations are not differentiable in the factors X_a and X_b, we have to use the second-order difference, rather then the second mixed partial derivative, to determine the contrast

$$\begin{aligned}
\Delta^2 E_{ab}(T_a + T_b; x_a, x_b) &= \Delta^2 E_a(T_a; x_a) + \Delta^2 E_a[E_b(T_b; x_b \mid T_a)] \\
&= 0 + \Delta^2 \int_0^\infty f_a(t_a; x_a)\, E_b(T_b; x_b \mid T_a = t_a)\, dt_a \\
&= 0 + \int_0^\infty [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})] \\
&\quad \cdot [E_b(T_b; x_{b2} \mid T_a = t_a) - E_b(T_b; x_{b1} \mid T_a = t_a)]\, dt_a \\
&= \int_0^{1.8} (e^{-t_a} - \tfrac{1}{4} e^{-(1/4)t_a})(\tfrac{1}{2} t_a^2 - \sqrt{2t_a})\, dt_a \\
&\quad + \int_{1.8}^\infty (e^{-t_a} - \tfrac{1}{4} e^{-(1/4)t_a})(-0.277)\, dt_a \\
&= -0.334368 + 0.130835 \\
&= -0.203533 < 0, \qquad (8)
\end{aligned}$$

indicating subadditivity.

(4) *Negatively dependent and subadditive.* For this case we can use the example provided at the end of the previous section discussing single factor reversals. Recall that

$$E_{ab}(T_a + T_b; x_a, x_b) = \frac{1}{x_a} + \frac{x_a}{x_b(x_a + 1)}.$$

In this model, subadditivity can be confirmed by observing that

$$\frac{\partial^2}{\partial x_a\, \partial x_b} E_{ab}(T_a + T_b; x_a, x_b) = -\frac{1}{x_b^2(x_a + 1)^2},$$

which is always negative.

(5) *Negatively dependent and additive.* Here, we employ a similar device used in the positively dependent and yet additive model presented earlier. Let there be, in addition to the serial subsystems S_a and S_b, a third subsystem, S_c, that affects the other two subsystems.

Define the probability functions as

$$f_a(t_a \mid c) = c x_a e^{-c x_a t_a}$$

$$f_b(t_b \mid c) = \frac{x_b}{c} e^{-(x_b/c) t_b}$$

$$\begin{aligned}
P_c(C) &= \frac{1}{2} \quad \text{for} \quad C = 2 \\
& \frac{1}{2} \quad \text{for} \quad C = 4.
\end{aligned}$$

Now, the total expected processing time becomes

$$\begin{aligned} E_{ab}(T_a+T_b; x_a, x_b) &= E_c[E_{ab}(T_a+T_b; x_a, x_b \mid C)] \\ &= E_c\left(\frac{1}{Cx_a}+\frac{C}{x_b}\right) \\ &= \frac{1}{x_a}E_c\left(\frac{1}{C}\right)+\frac{1}{x_b}E_c(C) \\ &= \frac{3}{8}\left(\frac{1}{x_a}\right)+3\left(\frac{1}{x_b}\right) \end{aligned}$$

for which $(\partial^2/\partial x_a\,\partial x_b)=0$. The negative dependence can be seen in the rates of S_a and S_b. When C is large, S_a tends to be faster while the rate of S_b has decreased.

(6) *Negatively dependent and superadditive.* Consider the following densities for S_a and S_b:

$$S_a : f_a(t_a; x_{a1}) = e^{-t_a}$$

$$f_a(t_a; x_{a2}) = 20e^{-20t_a}$$

$$S_b : f_b(t_b; x_b \mid t_a) = \frac{1}{R(t_a, x_{bi})}\, e^{(1/R(t_a, x_{bi}))t_b},$$

where $E_b(T_b; x_{bi} \mid T_a) = R(t_a, x_{bi})$ and $i=1, 2$. Now we construct $R(t_a, x_{bi})$ so that there exists a negative dependence between T_b and T_a. Define

$$E_b(T_b; x_{b1} \mid T_a = t_a) = \begin{cases} -t_a^2+2 & t_a \leqslant 0.925 \\ e^{-2t_a}+1 & t_a > 0.925 \end{cases}$$

and $E_b(T_b; x_{b2} \mid T_a = t_a) = e^{-2t_a}$. Note that for both levels of X_b, the expected processing time for S_b is a decreasing function of T_a reflecting the required negative dependence. To determine the predicted factorial interaction in this model we can obtain the second-order difference of the expected processing times for the four combinations of factor levels. In this case,

$$\begin{aligned} \Delta^2 E_{ab}(T_a+T_b; x_a, x_b) &= \Delta^2 E_a(T_a; x_a) + \Delta^2 E_a[E_b(T_b; x_b \mid T_a)] \\ &= 0 + \int_0^\infty [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})] \\ &\quad \cdot [E_b(T_b; x_{b2} \mid T_a = t_a) - E_b(T_b; x_{b1} \mid T_a = t_a)]\, dt_a \\ &= \int_0^{0.925} (20e^{-20t_a} - e^{-t_a})(e^{-2t_a} + t_a^2 - 2)\, dt_a \\ &\quad + \int_{0.925}^\infty (20e^{-20t_a} - e^{-t_a})(-1)\, dt_a \\ &= -0.325594 + 0.396531 \\ &= 0.070936 > 0. \end{aligned}$$

Because this second-order difference is positive, we can conclude that this model is superadditive.

Parallel

In the following models, all of which are based on exponential distributions, dependencies of total completion times are induced by changing the processing rates across stages. That is, when the first subsystem completes processing, the rate associated with the still operating subsystem will either increase, in the case of positive dependence, or it will decrease, in the case of negative dependence. This change of rate in the second stage may occur for both subsystems, or it may occur only for a given subsystem. Townsend (1974) and Townsend and Ashby (1983) showed that, for parallel exponential models, a faster second stage rate does, in fact, lead to positive dependence of total completion times, whereas a slower second stage rate will result in negative dependence of total completion times.

(1) *Positively dependent and superadditive.* Let $g_{ij}(t_j; x_i)$ be the parallel density of subsystem S_i, $i = a, b$, during stage $j = 1, 2$ and $X_a = X_b = \{1, 2\}$. Note that we are employing intercompletion times T_i on stage i rather than total completion time T. Then, the densities for the two subsystems in the four conditions are

Condition		S_a density	S_b Density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = 20e^{-20t_2} =$	$g_{b2}(t_2; 1)$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1}$,	$g_{b1}(t_1; 2) = 10e^{-10t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = 20e^{-20t_2}$,	$g_{b2}(t_2; 2) = 200e^{-200t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 10e^{-10t_1}$,	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 2:	$g_{a2}(t_2; 2) = 20e^{-200t_2}$,	$g_{b2}(t_2; 1) = 20e^{-20t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 10e^{-10t_1} =$	$g_{b1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = 200e^{-200t_2} =$	$g_{b2}(t_2; 2)$

If we denote $\langle a, b \rangle$ as the event that S_a finishes before S_b, then we can find

$$
\begin{aligned}
E(T; x_a, x_b) &= E_{ab}[\max(T_a, T_b); x_a, x_b] \\
&= P(\langle a, b \rangle)\, E(T_1 + T_2; x_a, x_b \mid \langle a, b \rangle) \\
&\quad + P(\langle b, a \rangle)\, E(T_1 + T_2; x_a, x_b \mid \langle b, a \rangle) \qquad (9)
\end{aligned}
$$

for each condition where T_1 and T_2 are the first and second stage intercompletion time random variables. $P(\langle a, b \rangle)$ is the ratio of the first stage rate of S_a to the sum of the first stage rates of S_a and S_b. For more details see Townsend (1974) and Townsend and Ashby (1983). Then, for each condition, from (9),

$$E(T_{11}) = \tfrac{1}{2} + \tfrac{1}{20} = \tfrac{11}{20},$$

$$E(T_{12}) = \tfrac{1}{11}(\tfrac{1}{11} + \tfrac{1}{200}) + \tfrac{10}{11}(\tfrac{1}{11} + \tfrac{1}{20}) = \tfrac{301}{2200} = E(T_{21}),$$

$$E(T_{22}) = \tfrac{1}{20} + \tfrac{1}{200} = \tfrac{11}{200}.$$

Now, the second-order difference, $\Delta^2 = E(T_{11}) - E(T_{12}) - E(T_{21}) + E(T_{22}) = 729/2200 > 0$. Therefore, this model predicts superadditivity.

(2) *Positively dependent and additive.* This model possesses a super-reallocation mechanism that renders the subsystems positively dependent. The densities for the four conditions are:

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = 2e^{-2t_2} =$	$g_{b2}(t_2; 1)$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1},$	$g_{b1}(t_1; 2) = 2e^{-2t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = 2e^{-2t_2},$	$g_{b2}(t_2; 2) = 4e^{-4t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1},$	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 2:	$g_{a2}(t_2; 2) = 4e^{-4t_2},$	$g_{b2}(t_2; 1) = 2e^{-2t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1} =$	$g_{b_1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = 4e^{-4t_2} =$	$g_{b2}(t_2; 2)$

The expectations are

$$E(T_{11}) = 1, \quad E(T_{12}) = E(T_{21}) = \tfrac{3}{4}, \quad E(T_{22}) = \tfrac{1}{2}.$$

Clearly, $\Delta^2 = 0$, so this model is additive.

(3) *Positively dependent and subadditive.* Here, the strength of the dependence is weak when compared to the magnitude of the factor effects. It is reasonable, then, that the factorial behavior is very much like an independent parallel model which produces subadditivity (see Townsend & Piotrowski, 1981; Townsend, 1984). The densities are

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = (8/7)\, e^{-(8/7)t_2} =$	$g_{b2}(t_2; 1)$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1}$,	$g_{b1}(t_1; 2) = 20e^{-20t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = (8/7)\, e^{-(8/7)t_2}$,	$g_{b2}(t_2; 2) = 21e^{-21t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 20e^{-20t_1}$,	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 2:	$g_{a2}(t_2; 2) = 21e^{-21t_2}$,	$g_{b2}(t_2; 1) = (8/7)\, e^{-(8/7)t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 20e^{-20t_1} =$	$g_{b1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = 21e^{-21t_2} =$	$g_{b2}(t_2; 2)$

The four expectations are

$$E(T_{11}) = \tfrac{11}{8}, \quad E(T_{12}) = E(T_{21}) = \tfrac{779}{882}, \quad E(T_{22}) = \tfrac{61}{840}.$$

The second-order difference, in this case, is $-703/2205 < 0$, which implies subadditivity.

(4) *Negatively dependent and subadditive.* In this model, the second stage rate for S_b is slower than that of its first stage rate, while the rate for S_a remains unchanged across stages. Therefore, it is a case where T_b depends negatively on T_a when S_a finishes first, rather than one of mutual dependence. This does not invalidate our results, because on the average T_b is a decreasing function of T_a. The densities for the four conditions are:

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = e^{-t_2}$,	$g_{b2}(t_2; 1) = (1/2)\, e^{-(1/2)t_2}$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1}$,	$g_{b1}(t_1; 2) = 2e^{-2t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = e^{-t_2}$,	$g_{b2}(t_2; 2) = e^{-t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1}$,	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 2:	$g_{a2}(t_2; 2) = 2e^{-2t_2}$,	$g_{b2}(t_2; 1) = (1/2)\, e^{-(1/2)t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1} =$	$g_{b1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = 2e^{-2t_2}$,	$g_{b2}(t_2; 2) = e^{-t_2}$

The expectations are then

$$E(T_{11}) = 2, \quad E(T_{12}) = \tfrac{4}{3},$$

$$E(T_{21}) = \tfrac{11}{6}, \quad E(T_{22}) = 1.$$

So, $\Delta^2 = -\frac{1}{6}$ which denotes subadditivity.

(5) *Negatively dependent and additive.* Here, we again introduce the mutual dependence produced by both subsystems slowing down in the second stage. The densities are given as

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = (3/4)\, e^{-(3/4)t_2} =$	$g_{b2}(t_2; 1)$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1}$,	$g_{b1}(t_1; 1) = 3e^{-3t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = (3/4)\, e^{-(3/4)t_2}$,	$g_{b2}(t_2; 2) = e^{-t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 3e^{-3t_1}$,	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 1:	$g_{a2}(t_2; 2) = e^{-t_2}$,	$g_{b2}(t_2; 1) = (3/4)\, e^{-(3/4)t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 3e^{-3t_1} =$	$g_{b1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = e^{-t_2} =$	$g_{b2}(t_2; 2)$

Here, the four expectations are

$$E(T_{11}) = \tfrac{11}{6}, \quad E(T_{12}) = E(T_{21}) = \tfrac{3}{2}, \quad E(T_{22}) = \tfrac{7}{6},$$

It is evident that $\Delta^2 = 0$, and thus this model is additive.

(6) *Negatively dependent and superadditive.* The only difference between this model and the previous one is that the effect of moving the factors to their higher levels is to increase the processing rate of only the first stage to 100. That is, $g_{a1}(t_1; 2) = g_{b1}(t_1; 2) = 100e^{-100t_1}$. All of the other densities are the same as before. The effect of this change is to render the negative dependence greater at the higher factor levels. Intuitively, this should make the factors superadditive. This is confirmed by looking at the resultant expectations:

$$E(T_{11}) = \tfrac{11}{6}, \quad E(T_{12}) = E(T_{21}) = \tfrac{406}{303}, \quad E(T_{22}) = \tfrac{201}{200}.$$

Now, $\Delta^2 = 3201/20200 > 0$ and, therefore, superadditivity results. ∎

The message in the above theorem brings up the possibility of obtaining another diagnostic of dependent systems. This diagnostic is the ability of either serial or parallel systems with nonselective influence to yield superadditivity for some levels of experimental factors and subadditivity for other levels of those factors. This change in contrast requires that there exist a dependence of some type, at least at the level of the expectations, in the system, but it can occur for either a positive or negative dependence in both serial and parallel models, as can be seen from the array of models contained in the proof of Theorem II. We state, for completeness, this result in the following theorem.

THEOREM III. *Assume the absence of direct nonselective influence. A change in contrast sign within a multilevel factorial experiment implies that there exists indirect nonselective influence in both serial and parallel systems.*

Proof of this theorem is achieved by recognizing that it is related to the contrapositive of two previously published propositions concerning the factorial behavior of independent serial and parallel models for exhaustive processing (see Townsend & Ashby, 1983, Chap. 12). One proposition stated that independence of processing in serial models combined with selective influence of the experimental factors produced additivity for all levels of the factors and all time, t. The other proposition associated independent parallel processing and selective influence of the factors with subadditivity. The independence assumption can be relaxed somewhat without altering the basic results, see footnote 3. Examining the contrapositive of each of these statements, we have that a failure to find additivity, assuming serial processing and the absence of direct nonselective influence (as always), requires a failure of stochastic independence of the processes (i.e., an indirect nonselective influence). Assuming parallel processing and the absence of direct nonselective influence, the failure to find subadditive factor effects implies the existence of an indirect nonselective influence. Note that the event of obtaining a change in contrast sign is a subset of the event of failing to find additivity for all levels of the factors in the serial case and failing to find subadditivity in the parallel case. Hence, truth of the original theorems concerning independence (or marginal selectivity) and factorial behavior entails the truth of Theorem III. Note that the models used in the proof of the previous theorem satisfy the question of the existence of dependent systems exhibiting contrast change for both positive and negative dependence. This is straighforward to observe in the case of the parallel models as it was the tradeoff between increases in the rates of processing for the two subsystems as the experimental factors were increased and the strength of the dependence between those subprocesses.

To review, Theorem I shows that a single factor reversal indicates that a negative dependence is present. Theorem II demonstrates that the type of dependence in a serial or parallel system is logically independent of the type of factorial interaction. Finally, Theorem III asserts that a change in contrast sign within an experiment requires the existence of a stochastic dependence.

It is natural to ask, at this time, is the presence of a reversal also unrelated to interaction type? Theorem IV demonstrates that the answer is yes. None of the negatively dependent models presented earlier, except for the serial subadditive case, contained a single factor reversal. Even in this excepted instance, there need not always be a reversal. Therefore, all that needs to be demonstrated is that reversals can occur for all three types of interaction. One additional fact that is apparant is that additivity is impossible if a reversal due to one factor occurs only for one level of the other factor. That is, if, say

$$E(T_{ab};x_{a1},x_{b1})-E(T_{ab};x_{a2},x_{b1})>0$$

and

$$E(T_{ab};x_{a1},x_{b2})-E(T_{ab};x_{a2},x_{b2})<0,$$

then $\Delta^2E(T_{ab};x_a,x_b)\neq 0$.

THEOREM IV. *Assume only indirect nonselective influence of the factors. Then, in both serial and parallel models, the presence of a single factor reversal in the mean* RTs *is not logically related to the type of factorial interaction produced in the system.*

It is straightforward to show that reversal can occur for both superadditivity and subadditivity by simply providing example models. However, in the case of additivity, it is more difficult to discover the appropriate combination of factor effects and negative dependence strength that leads to a zero interaction. Nevertheless, it is clear that by varying the relative strengths of the factorial influences and the dependence, any system that is continuous in these variables will pass through additivity when going from one contrast sign to the other.

Proof. As a proof we present the example models that yield a single factor reversal for both subadditivity and superadditivity, first in the serial case then in the parallel case. Note that, by Theorem I, these models must be negatively dependent.

(1) *Serial subadditive with reversal.* For this case, refer to the detailed example at the end of the discussion of Theorem I. Recall that $E_{ab}(T_a+T_b)=1/x_a+x_a/x_b(x_a+1)$ and $\partial^2/\partial x_a\,\partial x_b=-1/x_b^2(x_a+1)^2$, which is always negative, implying subadditivity. This model exhibits a reversal with respect to X_a whenever $(\partial/\partial x_a)\,E_{ab}(T_a+T_b;x_a,x_b)>0$, that is when $x_b<(x_a/(x_a+1))^2$. A specific set of numbers satisfying this condition was provided at the end of the discussion of the single factor reversal phenomenon.

(2) *Serial superadditive with reversal.* While somewhat contrived, this example satisfies our assumptions and does indicate the logical independence of the reversal phenomenon and contrast sign. Define the following densities and conditional expectations:

$$S_a: f_a(t_a;x_{a1})=e^{-t_a}$$

$$f_a(t_a;x_{a2})=10e^{-10t_a}$$

$$S_b: f_b(t_b;x_{bi}\mid t_a)=\frac{1}{R(t_a,x_{bi})}\,e^{(1/R(t_a,x_{bi}))\,t_b}.$$

So, $E_b(T_b;x_{bi}\mid T_a=t_a)=R(t_a,x_{bi}),\ i=1,2.$

Now, let

$$E_b(T_b,x_{b1}\mid t_a)=\begin{cases}-t_a^4+10 & 0<t_a<1.76515\\ 10e^{-2t_a} & t_a>1.76515\end{cases}$$

$$E_b(T_b;x_{b2}\mid x_{b2}\mid t_a)=9e^{-2t_a}.$$

The second-order difference, as defined in (8), is then

$$\begin{aligned}
&\Delta^2 E_{ab}(T_a + T_b; x_a, x_b) \\
&\quad = \Delta^2 E_a(T_a; x_a) + \Delta^2 E_a[E_b(T_b; x_b \mid T_a)] \\
&\quad = \Delta^2 E_a[E_b(T_b; x_b \mid T_a)] \\
&\quad = \int_0^\infty [f_a(t_a; x_{a2}) - f_a(t_a; x_{a1})] \\
&\qquad \cdot [E_b(T_b; x_{b2} \mid T_a = t_a) - E_b(T_b; x_{b_1} \mid T_a = t_a)]\, dt_a \\
&\quad = \int_0^{1.76515} (10e^{-10t_a} - e^{-t_a})(9e^{-2t_a} + t_a^4 - 10)\, dt_a \\
&\qquad + \int_{1.76515}^\infty (10e^{-10t_a} - e^{-t_a})(-e^{-2t_a})\, dt_a
\end{aligned}$$

$= 1.99126 + 0.00167145 = 1.99293 > 0$, which implies superadditivity. We have a single factor reversal whenever $\Delta_{xa} E_{ab}(T_a + T_b; x_a, x_b) > 0$. Examine this first-order difference for the level $X_b = x_{b1}$:

$$\begin{aligned}
\Delta_{x_a} &= \tfrac{1}{10} - 1 + \int_0^{1.76515} (10e^{-10t_a} - e^{-t_a})(-t_a^4 + 10)\, dt_a \\
&\quad + \int_{1.76515}^\infty (10e^{-10t_a} - e^{-t_a})(10e^{-2t_a})\, dt_a
\end{aligned}$$

$= -0.9 + 2.52378 - 0.0167145 = 1.60707 > 0$. Therefore, X_a reorders the means when $X_b = x_{b1}$. It can be shown that a single factor reversal also occurs when $X_b = x_{b2}$.

(3) *Parallel subadditive with reversal.* Define the densities as follows:

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1} =$	$g_{b1}(t_1; 1)$
	Stage 2:	$g_{a2}(t_2; 1) = e^{-t_2}$,	$g_{b2}(t_2; 1) = (1/8)\, e^{-(1/8)t_2}$
(1, 2)	Stage 1:	$g_{a1}(t_1; 1) = e^{-t_1}$,	$g_{b1}(t_1; 2) = 2e^{-2t_1}$
	Stage 2:	$g_{a2}(t_2; 1) = e^{-t_2}$,	$g_{b2}(t_2; 2) = (1/4)\, e^{-(1/4)t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1}$,	$g_{b1}(t_1; 1) = e^{-t_1}$
	Stage 2:	$g_{a2}(t_2; 2) = 2e^{-2t_2}$,	$g_{b2}(t_2; 1) = (1/8)\, e^{-(1/8)t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1; 2) = 2e^{-2t_1} =$	$g_{b1}(t_1; 2)$
	Stage 2:	$g_{a2}(t_2; 2) = 2e^{-2t_2}$,	$g_{b2}(t_2; 2) = (1/4)\, e^{-(1/4)t_2}$

The four expectations are $E(T_{11})=5$, $E(T_{12})=7/3$, $E(T_{21})=35/6$, and $E(T_{22})=5/2$. The second-order difference is $-2/3<0$, verifying subadditivity. We get a reversal for both levels of X_b because $E(T_{21})-E(T_{11})=5/6>0$ and $E(T_{22})-E(T_{12})=1/6>0$.

(4) *Parallel superadditive with reversal.* The densities for the four conditions are

Condition		S_a density	S_b density
(1, 1)	Stage 1:	$g_{a1}(t_1;1)=5e^{-5t_1}$,	$g_{b1}(t_1;1)=(1/10)\,e^{-(1/10)t_1}$
	Stage 2:	$g_{a2}(t_2;1)=5e^{-5t_2}$,	$g_{b2}(t_2;1)=(1/100)\,e^{-(1/100)t_2}$
(1, 2)	Stage 1:	$g_{a1}(t_1;1)=5e^{-5t_1}$,	$g_{b1}(t_1;2)=e^{-t_1}$
	Stage 2:	$g_{a2}(t_2;1)=5e^{-5t_2}$,	$g_{b2}(t_2;2)=(1/100)\,e^{-(1/100)t_2}$
(2, 1)	Stage 1:	$g_{a1}(t_1;2)=10e^{-10t_1}$,	$g_{b1}(t_1;1)=(1/10)\,e^{-(1/10)t_1}$
	Stage 2:	$g_{a2}(t_2;2)=10e^{-10t_2}$,	$g_{b2}(t_2;1)=(1/100)\,e^{-(1/100)t_2}$
(2, 2)	Stage 1:	$g_{a1}(t_1;2)=10e^{-10t_1}$,	$g_{b1}(t_1;2)=e^{-t_1}$
	Stage 2:	$g_{a2}(t_2;2)=10e^{-10t_2}$,	$g_{b2}(t_2;2)=(1/100)\,e^{-(1/100)t_2}$

Here, the four expectations are $E(T_{11})=25051/255\approx 98.2392$, $E(T_{12})=1253/15\approx 83.5333$, $E(T_{21})=100101/1010\approx 99.1099$, and $E(T_{22})=10011/110\approx 91.0091$. Therefore, the second-order difference is $124750/18887\approx 6.605>0$, confirming superadditivity. Note that $E(T_{21})-E(T_{11})\approx 0.8707>0$ and $E(T_{22})-E(T_{12})\approx 7.4758>0$. Thus, X_a reorders the means for both levels of X_b, and yet the system is superadditive.

Summary and Conclusion

A diagram that may be helpful in summarizing the results of serial and parallel models and their factorial behavior is shown in Fig. 3.

Beginning in the center and proceeding downward, we can note that if the subprocesses in the system are independent (or at least such that the dependence, if present, does not produce any indirect nonselective influence), serial models produce additivity as stated by Sternberg (1969) and proved under general conditions in Townsend and Ashby (1983) and Townsend (1984). Parallel models, on the other hand, with independence of processing exhibit subadditive factor effects (Schweickert, 1978; Townsend & Piotrowski, 1981; Townsend & Ashby, 1983).

Again, returning to the center but progressing upward in the figure, we can see that if a dependence exists between the subprocesses and it is strong enough, it may be detectable by using the results of the present paper. A single factor reversal can arise only in the presence of a negative dependence, but it need not always occur.

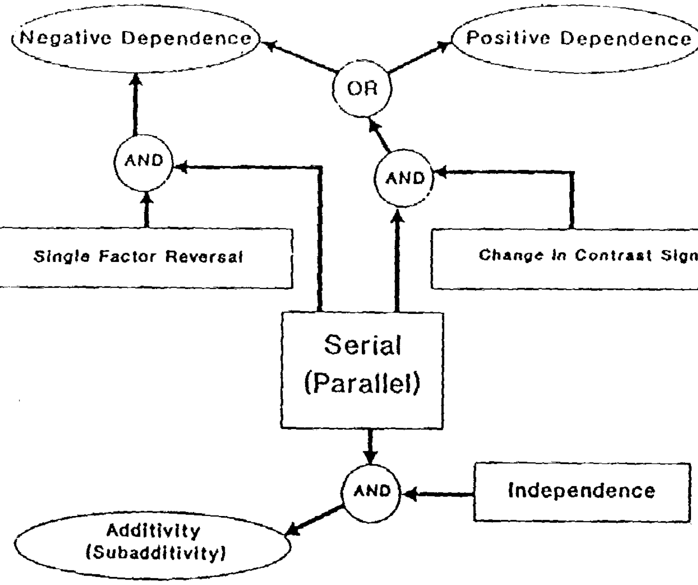

FIG. 3. An implication graph demonstrating the relationships between the various forms of dependencies and interactions. See text.

If the interactions change in sign (for example, start subadditive then become superadditive) across several factor levels, and the other assumptions are correct (i.e., parallel or serial, etc.), a dependence of some type must be present in the system. It would not be possible, however, to determine whether processing is serial or parallel in the dependent situation because these results apply equally to both cases.

There is an interesting aspect to note in this figure. The results concerning independent processing begin with the theoretical assumptions about the model structure and proceed to the experimental implications. Just the opposite is true for dependencies. The implications start with the empirical findings and point to the underlying processing structure that is necessarily responsible. This speaks to the complicating effects that stochastic dependencies are likely to create when drawing inferences from factorial interactions. The only underlying assumptions common to all of these results is that the experimental factors that are chosen indeed selectively influence their respective subsystems so that interactions are the consequence of the mental structure and its characteristics rather than a direct contamination by the inappropriate factors.

Can these results be fruitfully employed to help identify underlying architectures? In general, dependencies can clearly obfuscate that architecture by allowing a structure of one type to mimic the factorial contrast of another. However, single factor reversal and contrast change can, at least theoretically, aid in detecting the presence of dependencies. Positive dependencies never lead to factor reversal effects, but a sufficiently strong negative dependence generates a reversal. Conversely, if the potential architectures are confined to serial or parallel, then a reversal implies that a negative dependence exists, whichever system is present. Once a dependence of either variety is present, we have seen that it can be associated either with superadditivity, additivity, or subadditivity, so an experimentally determined reversal cannot be used to help separate parallel from serial processing. Contrast change, within the set of parallel and serial processes, and positing the required ancillary conditions, also implicates stochastic dependencies. It is too early to tell if these new features of processing can actually be effective in architecture identification in particular and model testing in general. One direction to go may be to construct experimental tasks that encourage strong negative dependencies and therefore possible reversals. Another is to perform factorial experiments over wide ranges of the factor levels to look for contrast change.

Some of the limitations in the application of these results have been alluded to, but they are worth discussing all together. It is possible that some other type of architecture and/or mode of operation is responsible for contrast change, rather than a dependent serial or parallel system. For instance , one of the other members of the stochastic latent network family (e.g., Schweickert, 1978, 1983a; Schweickert and Townsend, 1989) may be operating. Some converging evidence using other reaction time methods and accuracy measures could be used to marshal support for a serial or a parallel model that may be immune to the effects of dependencies among subsystems. A collection of such techniques is described in Townsend (1990a). Results from these, in conjunction with results from factorial methods, would give a fine grained account of the system of interest.

A word is also in order concerning the postulate that experimental factors always act monotonically, at least within the factor range of the experiment. With many factors, this assumption seems eminently reasonable, as gained from numerous experiments validating it, for example, brightness of a visual stimulus. If an experiment used an entirely new factor, of course, the assumption might be suspect until evidence was gained in its favor. One way of gaining this evidence is to construct a task that includes only one of the hypothesized subsystems from the larger system and manipulate the associated factor across multiple levels and perform a one-way ANOVA and determine if the effects are monotonic. As an example, suppose the investigator is interested in classification of visual stimuli. One of the constituent processes would presumably be sensory perception of the stimulus. To test if the factor, say brightness, has a monotonic effect on this perception, a simple reaction time task with the same stimuli could be utilized. This type of validation technique could be used in conjunction with any factorial experiment, not just those pertaining to stochastic dependencies.

A related question concerning factor effects is selectivity. We are, of course, allowing indirect nonselective influence through dependencies but we wish to rule out direct nonselective influence. Without assuming a priori the nature of the model (i.e., serial or parallel) it may be difficult to directly test this assumption. The usual scenario is to have accepted a particular architecture for a given task as a result of previous study, and to then test a new factor for its locus of effect in that architecture. For example, Miller and Pachella (1973) used standard additive factors logic in a memory scanning task to determine that variatons in stimulus probability influenced the stimulus encoding stage in a serial system. However, if one is not certain of the architecture, testing the selectivity of a new factor may be meaningless.

There are other theoretical considerations that may lead one to accept factor selectivity in the absence of a strong test. One such consideration could be due to anatomical characteristics. For example, Townsend and Nozawa (1988, 1992; see also Nozawa, 1992) found evidence for parallel processing of visually presented dots using factorial technology. In their paradigm, a dot stimulus was presented to one or the other of the subjects' eyes or to both, dichoptically. The two factors that were varied were the luminosity contrast of each dot. It is unlikely that the contrast of a dot in one eye would somehow affect the processing of the dot in the other eye. Hence, it is reasonable to assume there was direct selective influence of the factors. Unfortunately, not all processes of interest are so clearly physiologically delineated so that questions of selectivity may be of real concern. It should be noted, however, that, in practice, it is often assumed that experimental variables affect only the subsystems associated with them and that the validity of this assumption has not thwarted most of the research efforts (e.g., in latent mental network theory).

Although error data are not explicitly dealt with here, some remarks should be made. While no generally accepted theory exists for the relationship between reaction time and accuracy, Schweickert (1985) has developed some preliminary techniques for analysing both reaction time and accuracy within factorial designs. If each process is executed sequentially, then both mean reaction time and log percent correct would be additive in the factors (in addition to some ancillary assumptions). With respect to the theoretical developments we have outlined, we require that errors should not systematically vary with factor levels. Specifically, if a single factor reversal is present, it should not be due to a speed–accuracy tradeoff, nor should there be obvious and unusual patters of error data in change of contrast cases.

Another question is about future theoretical avenues. One obvious, if rather daunting, extension would be to more complex networks. Another concerns investigation of dependencies in more general distributional settings. In order to prove most of the present results, it sufficed to develop particular exponential or other specialized examples. It may be possible to develop general regularities involving dependence and systems factorial technology employing general dependence structures (e.g., Colonius, 1990).

ACKNOWLEDGMENTS

This research was supported in part by National Science Foundation Grant BNS 8710163, Memory and Cognitive Processes, and BNS 9112813, Human Cognition and Perception, to the first author. It was also aided by a NDSEG fellowship to the second author. The authors appreciate the helpful commentaries by the Action Editor, Roger Ratcliff, the Reviewer, Richard Schweickert, and an anonymous reviewer. The authors also thank Gabe Frommer for comments on an earlier draft.

References

Ashby, F. G. (1982). Testing the assumptions of exponential, additive reaction time models. *Memory and Cognition*, 10, 125–134.

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, 2, 93–123.

Colonius, H. (1990). Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*, 34, 253–275.

Donders, F. C. (1869). Die Schnelligkeit Psychisher Processe. *Archiv für Anatomie un Physiologie*, 657–681. (On the speed of mental processes. In W. G. Koster (Ed. and Transl.), *Attention and performance II*. Amsterdam: North-Holland. Reprinted from *Acta Psychologica*, 1969, 30.

Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-or-processing (OP) diagram. *Journal of Mathematical Psychology*, 27, 121–151.

Fisher, D. L., Sais, D., & Goldstein, W. M. (1985). Stochastic PERT networks: OP diagrams, critical paths, and the project completion time. *Computers and Operations Research*, 12, 471–482.

Goldstein, W. M., & Fisher, D. L. (1991). Stochastic Networks as models of cognition: Derivation of response time distributions using the order-of-processing method. *Journal of Mathematical Psychology*, 35, 214–241.

Gottsdanker, R., & Shragg, G. P. (1985). Verfication of Donders' subtraction method, *Journal of Experimental Psychology: Human Perception and Performance*, 11, 765–776.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be mislead: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like tast. *Memory & Cognition*, 7, 166–174.

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

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.

Miller, J., & Pachella, R. (1973). Locus of the stimulus probability effect. *Journal of Experimental Psychology*, 101, 227–231.

Nozawa, G. (1992). *Stochastic models of human information processing*. Ph.D. dissertation, Dartmouth College.

Posner, M. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.

Schweikert, R. (1977). *A generalization of the additive factor method*. (Tech. Rep. MMPP 77-2, Michigan Mathematical Psychology Program). Ann Arbor, MI: University of Michigan.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop tast. *Journal of Mathematical Psychology*, 18, 105–139.

Schweickert, R. (1980). Critical path sheduling of mental processes in a dual task. *Science*, 209, 704–706.

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

Schweickert, R. (1983a). Latent network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9, 253–383.

Schweickert, R. (1983b). Synthesizing partial order given comparability information: Partitive set and slack in critical path networks. *Journal of Mathematical Psychlogy*, 27, 261–276.

Schweickert, R. (1985). Separable effects of factors on speed and accuracy: Memory scanning, lexical decision, and choice tasks. *Psychological Bulletin*, 97, 530–546.

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

STERNBERG, S. (1966). High speed scanning in human memory. *Science*, 156, 652–654.

STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 176–315). Amsterdam: Elsevier/North-Holland.

THEIOS, J. (1973). Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7). New York: Academic Press.

TOWNSEND, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 10, 161–163.

TOWNSEND, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 25, 168–199.

TOWNSEND, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. New York: Erlbaum.

TOWNSEND, J. T. (1976a). A stochastic theory of matching processes. *Journal of Mathematical Psychology*, 14, 1–52.

TOWNSEND, J. T. (1976b). Serial and within-stage independent parallel model equivalence on the minimal completion time. *Journal of Mathematical Psychlogy*, 14, 219–238.

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

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

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

TOWNSEND, J. T., & ASHBY, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan and F. Restle (Eds.), *Cognitive Theory 3*. Hillsdale, NJ: Erlbaum.

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

TOWNSEND, J. T., & NOZAWA, G. (1988). Strong evidence for parallel processing with simple dot stimuli. Paper, presented at the *29th Meeting of the Psychonomic Society, Chicago*.

TOWNSEND, J. T., & NOZAWA, G. (1992). *On the spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories* (Research Report 65, Indiana University Cognitive Science Research Report Series). Bloomington, IN: Indiana University.

TOWNSEND, J. T., & PIOTROWSKY, S. (1981, Nov.). Parallel systems and factorial additivity of mean reaction time. Paper, presented at the *Meeting of the Psychonomics Society, Philadephia, PA*.

TOWNSEND, J. T., & SCHWEICKERT, R. (1985). Interactive effects of factors affecting processes in stochastic PERT networks. In G. d'Ydewalle (Ed.), *Cognition, information processing, and motivation*. Asterdam: 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, 309–327.

VORBERG, D. (1977). On the equivalence of parallel and serial models of information processing. Paper, presented at the *Tenth Annual Mathematical Psychology Meetings, University of California At Los Angeles*.

VORBERG, D. (1981). Reaction time distributions predicted by serial self-terminating models of memory search. In S. Grossberg (Ed.), *Proceedings of the Symposium in Applied Mathematics of the American Mathematical Society and the Society for Industrial and Applied Mathematics on Mathematical Psychology and Psychophysiology*. New York: Amer. Math. Soc.

Received: June 13, 1990

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium