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

1984 Vol. 28, No. 4, December 1984
Printed in Belgium

Uncovering Mental Processes with Factorial Experiments*

JAMES T. TOWNSEND

Purdue University

The analysis of reaction time (RT) additivity and interactions via factorial manipulation is a widely used and potentially powerful tool for elucidating mental processing; however, current implementation is limited by the scope of systems for which predictions are available, as well as statistical weakness. Predictions for an expanded set of system dimensions (serial/parallel, independent/dependent, and selective/nonselective influence) are given. The theorems are not limited to particular families of distributions. Statistical considerations are examined in the context of an exemplary nonadditive parallel system which adequately fits "additive" data but will only be rejected as additive when power is increased beyond the traditional criteria. © 1984 Academic Press, Inc.

INTRODUCTION

A major concern of many areas in modern psychology is the discovery of the internal processes responsible for performing various psychological tasks. Although the theory and methodology of this concern most properly may reside in the domain of cognitive psychology, nevertheless, social-personality, clinical, nonintrusive physiological, developmental, and educational psychology are all at one time or another interested in the delineation of mental processes that are potentially important in the behavior under study. This paper is about a popular and potentially powerful approach to such study and a beginning development of a firmer theoretical and methodological underpinning.

To establish some minimal terminology, we hereafter speak of the overall functioning "black box" as a "system" while "process" will be reserved for a unit or subsystem which performs some activity or task within the total system.[1] A

Address reprint requests to Dr. James T. Townsend, Department of Psychological Sciences, Purdue University, West Lafayette, Indiana 47907.

* Early reports of this study were given at the 1981 Society for Mathematical Psychology Meeting (Townsend, 1981a) and the 1981 Psychonomic Society Meeting (Townsend, 1981b). Much of the research reported here was supported by NSF Grant 0895-50. We thank Steven W. Link, Richard Schweickert, F. Gregory Ashby, and Ronald J. Evans for helpful comments on a draft of this paper.

[1] In some discussions it is useful to maintain a formal distinction between a "system," which represents an actual material entity functioning in real time and space and an abstract "model" of the system, usually specified insofar as possible in mathematical language (e.g., Townsend and Ashby, 1983). However, it will be less cumbersome to use a common terminology here for the two as there seems to be no danger of confusion. When we emphasize the properties of a set of processes we will usually speak of a "system," whereas when, particularly in the later sections, statistical aspects are a point of focus, "model" will be sometimes preferable.

0022-2496/84 $3.00

"subprocess" will indicate a further breakdown of a process into micro-level activities taking place within a process.

For instance, one may investigate the overall system engaged in reading a brief display of letters and reporting these to an experimenter. An example of a process within the system might be the subsystem responsible for recognizing individual letters (e.g., Townsend, 1981c). A possible subprocess of this particular process might be a feature identification mechanism operating on a within-letter basis.

Since the nineteenth century, an important approach has been oriented toward the study of reaction time (RT) and considerable theoretical and experimental literature and methodology have developed in support of this approach.

A historically critical innovation came with the "method of subtraction" associated primarily with the nineteenth century physiologist Donders (1969) but employed and further developed by W. Wundt and J. M. Catell among others. Essentially the method assumes that the mental processes under study take place in a nonoverlapping, successive fashion. This is an assumption of *seriality*.[2] Secondly, it is assumed that a process may be added or deleted from the total train of processes by manipulating the psychological task to be performed. This is the assumption of *pure insertion* and it is implicitly taken for granted that processes other than the one of interest are not affected by the experimental manipulation. In its simplest form, the average RT to a single stimulus, single response situation (*simple RT*) was subtracted from the average RT involving a choice from two alternatives in a 2-stimulus, 2-response situation where each stimulus was paired with a unique response (*choice RT*). The difference in average RTs was taken as the duration required for the cognitive aspects, choice, and discrimination of the process.

Despite the fact that the method has come under a good deal of criticism and obviously is based on very strong postulates, it continues to be employed from time to time. When the hypothetical processes are likely to be substantially separated in the processing chain and the pure insertion assumption does not seem too implausible, it may permit sound results. Ashby and Townsend (1980) and Ashby (1982) have recently suggested more powerful methods of testing pure insertion and related assumptions. As Kuhn (1962) and others have suggested, a scientific paradigm or methodology, even if far from perfect, usually survives until supplanted by something better. This is probably one good reason for the relatively long life of the method of subtraction, even employing the older, weaker tests.

Little over a decade ago, a significant bid for the "something better" title was put forth by Sternberg (1969a, 1969b). The *additive factor method*, as it has come to be known, effectively weakens the assumptions standing behind the method of subtraction thus broadening the sphere of theoretical systems amenable to investigation. The method is based on serially arranged processes, as is the

[2] The term "stage" has often been used to designate a process. In the past, "stage" has denoted or tended to connote seriality. Also, we employ the word "stage" to represent an intercompletion time (see below or Townsend, 1974) which may or may not be associated with a single process. For these reasons, we will as noted above use the less loaded terms "process" or "subsystem" even in strictly serial systems.

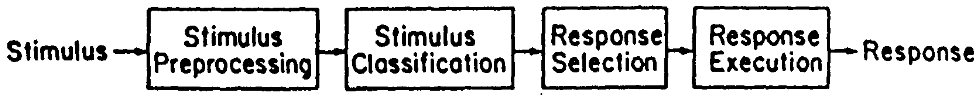

FIG. 1. The processing model of choice RT proposed by Smith (1968).

Donderian approach. However, if the subprocesses are indeed serially arranged and if (a) experimental factors (the traditional "independent variables") can be found that separately affect distinct processes, and (b) the successive processing times are stochastically independent, then the overall average RT will be an additive function of the experimental factors. (Note, however, that mean RT additivity is possible (but not implied) without stochastic independence of processing times.)

The postulation of separate effects by two or more experimental factors is known as the "assumption of selective influence." In fact, it was Sternberg's proposal that two hypothetical processes be defined as distinct only if two such additive experimental factors could be discovered. If the average RT was found to be an interactive function of the experimental factors, then, either two separate psychological processes of the supposed type did not exist, or the appropriate experimental factors had not yet been found. Yet again, if two (or more) factors interacted then they were said to affect the same process. Therefore, the possibility that a single process could yield an additive mean RT function of two or more experimental factors was excluded.[3] The case where two or more separate processes might interact was also excluded or perhaps defined as a single process.

The prominence of the information-processing paradigm has led to the use of additive factors logic in diverse areas of psychology. The primary focus is, of course, in the cognitive literature wherein the models proposed are most often of the form of Smith (1968) or Sternberg (1969a). Smith hypothesized four serially arranged processes: stimulus preprocessing, stimulus classification, response selection, and response execution. See Fig. 1. In light of such a conceptualization, examples of factors used are stimulus quality, semantic context, size of positive set, stimulus probability, response type, response discriminability, and type of match.

The additive factor method as well as generalizations and extensions (e.g., Schweickert, 1978; Taylor, 1976; McClelland, 1979; Ashby and Townsend, 1980; Ashby, 1982; Pieters, 1983) are of potential application to any discipline interested in elucidating mental functioning. Thus, in addition to the multitude of studies done within the human information processing area proper (a number of which are mentioned below) the additive factor method has been employed to a greater or lesser extent in most other areas of psychology. Thus, Simon and Pouraghabagher (1978) employed the technique in order to specify the processing locus of particular developmental effects. Rogers (1974) investigated personality functions within the paradigm. Hemsley (1976) used the method to ascertain the source of the processing deficit in schizophrenics. P300 latencies, as well as RTs, were investigated in a factorial experiment by Gomer, Spicuzza, and O'Donnell (1976).

[3] Schweickert (1983) gives one of the few plausible examples of such a case of which we are aware.

Despite the importance and wide applicability of the additive factor method and related techniques, their effectiveness in producing clear-cut results that push us to an incontrovertibly higher understanding of human cognition is limited. Even the two basic postulates mentioned above can be questioned, insofar as (a) a single process could be additively affected by two factors and (b) two factors affecting separate processes could interact. These possibilities will be discussed further (see also, Taylor, 1976; Townsend and Ashby, 1983).

There are many statistical and philosophical pitfalls associated with the additive factor method as it is ordinarily practiced. Only certain of these which are integrally related to the specific thrust of this paper can be dealt with here, but the reader is referred to Theios (1973), Pachella (1974), Taylor (1976), Ashby and Townsend (1980), Pieters (1983) and Townsend and Ashby (1983) for further discussion.

Our major concern here is with the uncertainty which clouds theoretical conclusions drawn from additive factor results. We know that seriality and selective influence are sufficient to produce factorial additivity, yet it is not clear (a) just what other types of arrangements of processes might also predict factorial additivity or (b) what types of arrangements may contribute some plausibility to certain types of serial processes. A finding of additivity may also be implicitly augmenting the support for other nonserial systems. Conversely, we know very little about what varieties of subprocesses are ruled out by a discovery of additivity or what conclusions should be drawn in the event of strong interactions. Another dimension of here is that, depending on the salience of interactions produced by nonadditive systems, such nonadditive activity may be masked inadvertently by experimental noise and inappropriate conclusions drawn. Moreover, although a number of other types of processes intuitively might produce interactions, there has been almost no mathematical work in the litrature demonstrating such conclusions rigorously.[4]

This type of concern grows naturally out of an increased attention to problems connected with empirically testing models of distinct kinds of systems which may yield indistinguishable predictions (see, e.g., Townsend, 1974, 1972; Anderson, 1976). Such questions are a healthy sign in a maturing science.

It may be asked as to how the results reported below fit into the global problem of parallel–serial equivalence and distinguishability given the findings of severe problems in a review of empirical situations (Townsend, 1972, 1974, 1976a, 1976b). It turns out that the investigation of factorial additivity is one pathway whereby it is at least possible in principle to discriminate reasonably large classes of parallel and serial systems (for other potential methods, see Townsend and Ashby, 1983). Moreover, we shall see that a special case, where parallel models predict mean RT additivity, forms one of the earlier discovered instances where parallel models are mathematically equivalent to their serial counterparts.

Our approach, as espoused, for example, in Townsend and Ashby (1983), is based

[4] As exceptions, see Theios, Smith, Haviland, Traupmann, and Moy (1973), Taylor (1976), Schweickert (1978), McClelland (1979), and Ashby and Townsend (1980). The results differ with regard to the classes of systems treated, degree of mathematical rigor (e.g., some are based on computer simulation rather than closed mathematical proof), and depth of treatment.

on the foregoing viewpoint. Although we view the test for additivity as one important strategy in an overall systemic approach to uncovering psychological processes, we do not identify processes with additivity. Rather, we believe that processes should be studied in terms of their relationships, whether they function in series or simultaneously (i.e., in *parallel*) or in some other hybrid fashion, one significant aspect being the implications borne by these relationships for RT as affected by pertinent experimental factors. This study is an attempt to gather together (and develop when necessary) the predictions made by two major opposing types of systems: those positing a serial arrangement of processes as above and those based on a set of parallel (simultaneous) processes. In addition, certain strategic statistical questions and an overview of some related methodological problems we found in surveying the rather extensive literature involving the additive factor method will be taken up. The reader who wishes to gain more intuitive flavor of the empirical literature and statistical problems may wish to skim the section Applying Factorial Reaction Time Methods before proceeding.

It is important to note at the outset that the following theorems (Propositions 1–6) are distribution free. That is, the classes of probability laws covered and the proofs are not limited to any specific family (e.g., exponential, gamma, Weibull, etc.). Particular distributions are, of course, used to illustrate the propositions and to explicate other concepts.

Serial vs. Parallel Systems and Their Factorial Predictions

Before exhibiting the basic results having to do with these two prototypical types of processing we need to introduce some terminology. Let S_a and S_b refer to two processes (i.e., subsystems) involved in some type of psychological processing. Let T_a be the random variable representing the processing time of process S_a and define T_b similarly for S_b. Thus, probability density functions (see, e.g., Parzen, 1960) will be used for the development of the results concerning processing times. The residual time contributed by extraneous processes will be denoted T_0, which we will usually be able to ignore.

A factorial experiment can be identified in terms of three dimensions, as given in Fig. 2. The first classifies whether the subsystems operate serially or in parallel. If two subsystems S_a and S_b function serially then S_a operates on its input first, finishes, and then S_b begins processing, with no overlap in processing time. It is also typically assumed that the time lag between the output of S_a (and instant of completion) and the moment that S_b begins processing is nil, but this restriction may be relaxed with no harm to the predictions. The total processing time is, of course, the sum of the two components plus that attributed to processing components external to S_a and S_b, which we have denoted S_0.

Processing might also be parallel, in which event both S_a and S_b begin processing simultaneously but may finish at different times. The total processing time is simply the maximum time that either takes to complete its job.

The stochastic independence vs. dependence of the actual processing times is the

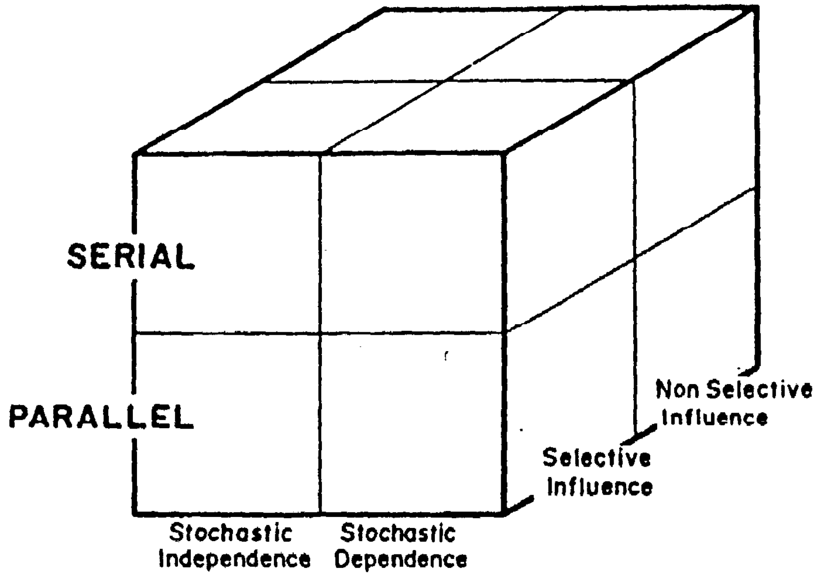

FIG. 2. The dimensions of interest for any factorial reaction-time experiment.

second dimension of interest, wherein independence means that knowledge of the processing time for one subsystem does not provide information regarding the processing time of the other. Note, however, that in serial systems the actual processing times under consideration are intercompletion times; while, in parallel systems, they are completion times.

The final facet, selective vs. nonselective influence, refers to the characteristics of the factor effects on the processing system. It will be convenient to assume that the experimental factors x_a and x_b are represented on a numerical scale so that, say, x_a is a value in some appropriate unit and the usual ordering and other properties of the number system hold in the separate scales. It will not be necessary here to represent the units explicitly.

Selective influence holds if factor x_a affects only subsystem S_a and if factor x_b affects only subsystem S_b. A unique processing locus is thus defined for a given factor effect. Nonselective influence occurs when a given process is influenced by more than one of the experimentally manipulated variables.

The following discussion on the results obtainable by factorial reaction-time experiments on processing systems will be divided into serial and parallel systems. Under each, the stochastic independence vs. dependence and selective vs. nonselective influence questions will be considered. Another conceptualization of these possibilities and their predicted results are given in Table 1. (Further explanation of this table will be presented later.) Proofs of the propositions given in the following sections can be found in Townsend and Ashby (1983, Chap. 12).

Basic Results on Serial Systems

In a serial system, the total processing time is the sum of the component processing times for S_a, S_b, and S_0. Thus, let $RT = T_a + T_b + T_0$. A specific value of RT will

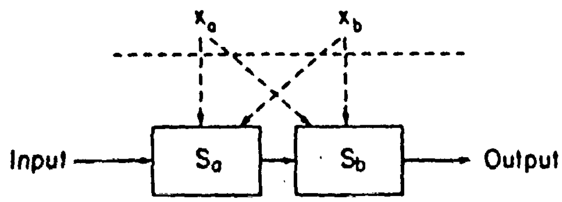

FIG. 3. The prototypical serial arrangement of processes S_a and S_b, including potential nonselective influences of the factors x_a and x_b.

be $rt = t_a + t_b + t_0$. Figure 3 shows the prototypical serial arrangement of processes S_a and S_b including potential nonselective influences of the x_a and x_b factors. Call the sum $T_a + T_b$ the overall processing time T. Let the density function for process S_i ($i = a,b$) under the influence of particular values of factors x_a and x_b be written as $f_i(t_i; x_a, x_b)$. Note that we have to allow for the possibility that selective influence fails, i.e., that either factor influences one or more of the density functions.

Now, because S_a operates first, the time of S_b cannot in a real system directly affect T_a. However, it may be that whether T_a is long or short may affect how long T_b is so we require the notation of conditional probability density function in the case of S_b, namely, $f_b(t_b; x_a, x_b \mid t_a)$, where t_a is the duration consumed by S_a on the current trial. Note that T_b is independent of T_a when $f_b(t_b; x_a, x_b \mid t_a) = f_b(t_b; x_a, x_b)$ for all t_a and t_b. Finally, we shall assume that T_0 is in serial arrangement with the two processes of interest and moreover is stochastically independent of them.

Next, we must write the expression for the overall mean or expectation of the RT:

$$E(\mathrm{RT}; x_a, x_b) = \int_0^\infty \int_0^\infty \int_0^\infty (t_a + t_b + t_0) f_a(t_a; x_a, x_b) \times f_b(t_b; x_a, x_b \mid t_a) f_0(t_0)\, dt_a\, dt_b\, dt_0. \quad (1)$$

This can easily be rewritten as

$$E(\mathrm{RT}; x_a, x_b) = E(T_a; x_a, x_b) + E_a[E_{b|a}(T_b; x_a, x_b \mid T_a)] + E(T_0). \quad (2)$$

The first and third Es on the right-hand side of Eq. (2) are written without subscript because it is taken for granted that the expectation is with regard to the t_a and t_0 densities, respectively. In the case of t_b, however, it is required that we first integrate with respect to t_b for a given t_a and then integrate over all t_a, as is indicated in the expectation operator notation. From Eq. (2), it is evident that T_0 is without effect with regard to the influence of the experimental factors so may be disregarded. More importantly, notice that potentially x_a might influence S_b through the stochastic dependency of t_b on t_a inherent in $f_b(t_b; x_a, x_b \mid t_a)$ even if f_b were not directly dependent on x_a.

This brings up the notion of two types of violations of selective influence. The first occurs when more than one experimental factor directly affect a given subsystem and is written into the density function. Thus, the density $f_a(t_a; x_a, x_b)$ identifies a "direct" nonselective influence in which S_a is affected by factors x_a and x_b. Note that both a stochastically dependent and independent system may exhibit direct nonselective influence. The second type of nonselective influence is "indirect," wherein x_a may affect S_b through the influence of values of T_a on T_b. That is, the stochastic dependency of processing times induces an influence of an otherwise unassociated factor x_a on a subsequent process S_b. This type occurs only for stochastically dependent systems.

Obviously, in serial systems, when T_b is independent of values of T_a, any influence must be direct. In dependent systems, however, three types of nonselective influence may occur—direct, indirect, or both. As examples, consider the following exponential density functions for S_b, where the factor effects are manifested in the rate parameters:

(1) direct: $f_b(t_b; x_a, x_b) = x_a x_b e^{-x_a x_b t_b}$, $t_b > 0$;

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

(3) direct and indirect: $f_b(t_b; x_a, x_b \mid t_a) = (x_a x_b)/t_a e^{-(x_a x_b t_b)/t_a}$, $t_a > 0$, $t_b > 0$.

In the direct case, the density of t_b is a function of both factors x_a and x_b, since the rate parameter is the product of these two factors. While the indirect case does not have x_a written in its density, because $f(t_a)$ is a function of x_a, an indirect nonselective influence is present. The final case shows a combination of direct and indirect influences.

Thus, note that when there is stochastic dependence the situation is more ambiguous in the sense that a given influence of x_a on T_b may sometimes be interpreted as direct or indirect or both through values of T_a. One of the interpretations may be more natural. In addition, it is even possible that a system with both direct and indirect nonselective influence could yield a distribution which shows selective influence. We cannot go into detail on this direct vs. indirect influence ambiguity here, but it will surface again in the discussion of parallel models. Typically, we simply accept the most obvious or natural interpretation.

Clearly, it is important that the x_a and x_b factors have a nontrivial influence on S_a and S_b, respectively. Often it will be the case that $E(T_i; x_i)$ will be a monotonic function of x_i, decreasing if x_i speeds up processing, increasing otherwise.

Another point concerns the situation where T_a, say, is itself an additive function of x_a and x_b. We typically assume for convenience of expression that this does not happen, but its presence would in no way affect our conclusions regarding mean RT additivity and parallelity vs. seriality. A serial system will produce factorial additivity whether, say, x_a affects two processes S_{a1} and S_{a2} or just one S_a (as long as it does not affect S_b). Similarly, the parallel systems below which fail to yield mean RT additivity do so whether or not one or more subprocesses make up each subprocess.

We are now in a position to state major results about serial mean RT additivity. The first two theorems give necessary and sufficient conditions for additivity to hold.[5]

PROPOSITION 1. *A serial system is mean RT additive if and only if the processes S_a and S_b are selectively influenced by x_a and x_b, respectively. In subsystem S_a, this*

[5] Proofs of the propositions are given in Townsend and Ashby (1983). Much of the explication and several examples as well as the statistical treatment coming later are new in this paper.

amounts to excluding direct nonselective influence, while in S_b both direct and indirect nonselective influence must be ruled out. We would then have

$$E(\mathrm{RT}; x_a, x_b) = E(T_a; x_a) + E(T_b; x_b) + E(T_0)$$

for all x_a and x_b.

This is clearly an additive function of x_a and x_b. Note that although it is possible to have stochastic dependency of the S_b density f_b on t_a and yet satisfy additivity, it is by no means guaranteed and in fact in most cases would probably not occur.

The next result gives a well-known sufficient condition for additivity in serial systems; it is essentially a corollary of Proposition 1.

PROPOSITION 2. *A serial system is mean* RT *additive if S_a and S_b are stochastically independent and selectively influenced by their respective factors x_a and x_b. Here it is sufficient that S_b is not directly nonselectively influenced by x_a.*

We remark that if T_a and T_b are stochastically dependent, selective influence holds on S_a, and S_b is not directly influenced by S_a, then S_a contributes an additive component to the overall mean RT. Thus, any nonadditive component would then come completely from $E_a[E_{b|a}(T_b; x_b \mid T_a)]$.

PROPOSITION 3. *If selective influence fails directly or indirectly (i.e., because of a stochastic dependency between S_a and S_b), then mean* RT *nonadditivity is possible.*

That direct nonselective influence can lead to mean RT nonadditivity is easily seen, but it is of interest to show how stochastic dependence, which yields an indirect nonselective influence, can also produce an interactive mean RT function of x_a and x_b.

A Nonadditive Serial System

As an example of a stochastically dependent serial system which yields a mean RT interaction, consider the exponential density functions

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

In this system, S_b is positively dependent on S_a in the sense that when t_a is large, t_b tends to be large. To begin with, simply note that the factor x_a is the rate of the exponential density of S_a. Large x's imply fast times and vice versa. Second, observe that in a sense the rate for S_b *conditional* on a specific time t_a may be viewed as x_b/t_a, i.e., the *conditional* rate is smaller for longer t_a values. Thus, we can expect slow S_a trials to be associated with slow S_b trials.

One way of proving this claim is to calculate the conditional cumulative distribution on t_b, namely

$$F_b(t_b; x_b \mid t_a) = 1 - e^{-(x_b t_b)/t_a},$$

where we see that F_b is a monotonic decreasing function of t_a. That is, for any t_b, increasing t_a always makes it less likely that S_b has completed processing by that time t_b. It follows that the conditional expected mean is a monotonic increasing function of t_a (the average time to finish in S_b increases the larger is t_a). That is, $E_{b|a}(T_b; x_b \mid t_a) = t_a/x_b$ which carries the obvious implication.

Finally, the overall contribution of S_b, averaged across all values of T_a is $E_a[E_{b|a}(T_b; x_b \mid T_a)] = 1/(x_a x_b)$. The total mean time, engaging Eqs. (1) and (2), is then just $E(T) = 1/x_a + 1/(x_a x_b)$.

We now compute a difference of theoretical mean RTs for a 2×2 factorial experiment which will be zero if and only if mean RT additivity holds. This is,

$$E(T; x_{a1}, x_{b1}) - E(T; x_{a1}, x_{b2}) - [E(T; x_{a2}, x_{b1}) - E(T; x_{a2}, x_{b2})],$$

where, of course, the expectations are taken over both T_a and T_b and where $x_{a2} > x_{a1}$, and $x_{b2} > x_{b1}$, denoting the two levels of each factor. The present serial dependent model will reveal *superadditivity* because the value of the expression will be greater than zero. In other words, the mean RT is elongated even more by a decrease in x_b when x_a is small than when it is larger.

Basic Results on Parallel Systems

In parallel systems, the overall processing time T is not the sum of the two processing times of S_a and S_b as it was in the serial system, rather it is the maximum or longest processing time associated with the two processes, $\max(T_a, T_b) = T$. As viewed in Fig. 4, the processes commence at the same instant and we leave open the possibility that the experimental factors bear nonselective influences on the two processes. The T_0 component will still be added to T and the other assumptions about its behavior are in force.

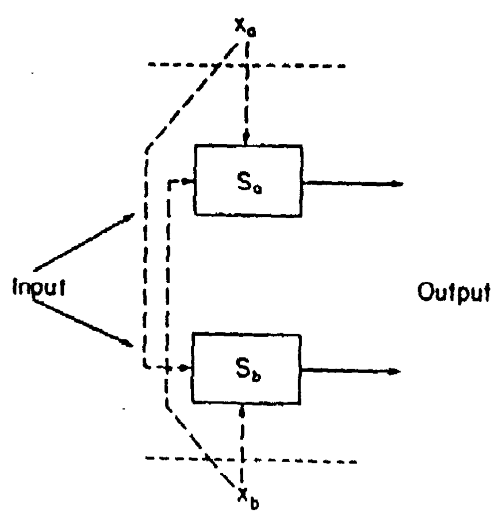

FIG. 4. The prototypical parallel arrangement of processes S_a and S_b, including potential nonselective influences of the factors x_a and x_b.

We may write the joint density function for the general parallel system as $g_{ab}(\tau_a, \tau_b; x_a, x_b)$, although we will usually be able to work with the joint cumulative distribution function or the distribution function on T itself. The former is

$$P_{ab}(T_a \leqslant \tau_a, T_b \leqslant \tau_b; x_a, x_b) = G_{ab}(\tau_a, \tau_b; x_a, x_b)$$
$$= \int_{\tau_a=0}^{\tau_a} \int_{\tau_b=0}^{\tau_b} g_{ab}(\tau'_a, \tau'_b; x_a, x_b)\, d\tau'_b\, d\tau'_a, \tag{3}$$

while the distribution on T is

$$P_{ab}(\max(T_a, T_b) \leqslant t; x_a, x_b) = P_{ab}(T \leqslant t; x_a, x_b)$$
$$= G_{ab}(\tau_a, \tau_b; x_a, x_b)\big|_{\substack{\tau_a=t \\ \tau_b=t}}. \tag{4}$$

That is, it is simply the joint distribution function evaluated at $\tau_a = \tau_b = t$.

In general, nonselective influences might arise either directly from an influence of an experimental factor on the alternative process or indirectly from a possible stochastic dependency between T_a and T_b. A special case of considerable interest in the cognitive and perceptual literature occurs when T_a and T_b are stochastically independent. The concept of independence is used in several different ways in the literature and some of these are sometimes confused with one another. The relations and contrasts among these are discussed with regard to RT models in Townsend (1974) and regarding psychophysics in an upcoming paper (Ashby and Townsend, submitted for publication). Here, the appropriate definition of stochastic independence of processing times in parallel systems is

$$G_{ab}(\tau_a, \tau_b; x_a, x_b) = G_a(\tau_a; x_a, x_b)\, G_b(\tau_b; x_a, x_b).$$

That is, the probability that S_a finishes at or before τ_a and S_b finishes at or before τ_b can be expressed as a product of the separate probabilities. Observe that the potential for direct selective influence remains open but the opportunity for indirect nonselective influence is closed.

The next task is to define the expectation of RT and T. It is clear that one must integrate the joint density over the parts of the probability space where τ_a is larger, times τ_a, and add that to the integral over the parts where τ_b is larger, times τ_b,

$$\begin{aligned}
E(RT) &= E(T + T_0; x_a, x_b) \\
&= \int_{t_0=0}^{\infty} \int_{\tau_b=0}^{\infty} \int_{\tau_a=\tau_b}^{\infty} (\tau_a + t_0)\, g_{ab}(\tau_a, \tau_b; x_a, x_b) f_0(t_0)\, d\tau_a\, d\tau_b\, dt_0 \\
&\quad + \int_{t_0=0}^{\infty} \int_{\tau_a=0}^{\infty} \int_{\tau_b=\tau_a}^{\infty} (\tau_b + t_0)\, g_{ab}(\tau_a; \tau_b; x_a, x_b) f_0(t_0)\, d\tau_b\, d\tau_a\, dt_0 \\
&= E(T; x_a, x_b) + E(T_0).
\end{aligned} \tag{5}$$

However, the expression involving T may be written in a way that makes Proposition 4 easy to prove. Although the proof will not be provided here (see Townsend and Ashby, 1983, p. 170), it may aid the reader's intuition to see the appropriate formulation. It is

$$E(T; x_a, x_b) = \int_{t=0}^{\infty} [1 - G_{ab}(t; x_a, x_b)]\, dt. \tag{6}$$

Thus, the expectation or average of T can be expressed as the integral of one minus the distribution function on the maximum completion time of the two times T_a and T_b. In computing the mean RT one would simply add $E(T_0)$.

Pursuant to achieving our result, we assume for convenience (it is not necessary) that both x_a and x_b are simply ordered, that is, the objects of each may be ordered relative to one another by a relation that is transitive, connected, and antisymmetric. This is, compatible with our idea that the factors may be expressed in terms of some numerical scale. Obviously, T_0 plays no role in mean RT additivity under our earlier assumptions so may henceforth be neglected. If the overall time T is to be an additive function of x_a and x_b then for any set of pairs of values (x_{a1}, x_{a2}) and (x_{b1}, x_{b2}), where $x_{a1} < x_{a2}$ and $x_{b1} < x_{b2}$, it must be the case that

$$E(T; x_{a1}, x_{b1}) - E(T; x_{a1}, x_{b2}) = E(T; x_{a2}, x_{b1}) - E(T; x_{a2}, x_{b2}). \tag{7}$$

Equation (7) represents a simple prototypical experiment where two levels of each factor are chosen and indicates that the effects on the average overall processing time are additive. (This is equivalent to the difference of theoretical means used earlier in our development of a nonadditive serial system.) With the above development in hand, including Eqs. (3)–(7), it becomes feasible to demonstrate Proposition 4.

Proposition 4. (A) *A sufficient condition to render mean RT additivity impossible for given fixed pairs (x_{a1}, x_{a2}) and (x_{b1}, x_{b2}) is that the mixed second-order difference in the distribution function on* $\max(T_a, T_b)$ *satisfy the constraint*

$$G_{ab}(t; x_{a1}, x_{b2}) - G_{ab}(t; x_{a1}, x_{b1}) - [G_{ab}(t; x_{a2}, x_{b2}) - G_{ab}(t; x_{a2}, x_{b1})]$$
$$>0 \qquad \text{for all } t \geqslant 0,$$

or

$$<0 \qquad \text{for all } t \geqslant 0. \tag{8}$$

That is, if the difference in the distribution function on the maximum processing time for two different levels of the same factor when the other factor is held constant is always larger or always smaller than that difference when the fixed factor is larger, then additivity is impossible.

(B) *If the direction of the order is > then superadditivity results, whereas if it is < then subadditivity occurs.*

Observe that the "always" in Proposition 4(A) is in terms of time t, not in terms of the experimental factors which are fixed for the given experiment. If one next supposes that the condition of Proposition 4 is in force for *all* values in x_a and x_b then a mean RT interaction will occur in all experiments (ignoring experimental and statistical error) involving different factor levels. The next proposition, in essence a corollary of the previous one, develops the important case of the independent parallel system.

PROPOSITION 5. (A) *If the following conditions hold for the distribution functions on the* $\max(T_a, T_b)$, *in terms of Eq.* (8) *in Proposition* 4(A), *then the condition of Proposition* 4 *is satisfied and mean RT additivity is impossible.*

(i) *Stochastic independence of processing times: For all values of* τ_a *and* τ_b,

$$G_{ab}(\tau_a, \tau_b; x_a, x_b) = G_a(\tau_a; x_a)\, G_b(\tau_b; x_b);$$

(ii) *Selective influence holds on the pertinent distributions of* (i): *For all values of* τ_i,

$$G_i(\tau_i; x_a, x_b) = G_i(\tau_i; x_i), \qquad i = a, b;$$

(iii) *The distribution functions derived from* (i) *and* (ii) *are strict monotonic increasing or decreasing functions of their respective factors*;

(B) *Further, if the distribution functions are written in such a way that the monotonicity goes in the same direction for each, that is, both are increasing or both are decreasing, then subadditivity always occurs.*

The idea that independent parallel models might predict subadditivity was suggested by Sternberg (1969a). As far as we know, the mathematical conditions for its attainment have never been explicitly stated.

When x_a and x_b lie on a continuum and the distribution functions are differentiable in their respective factors then (iii) of Proposition 5 can be stated as

$$dG_i(t_i; x_i)/dx_i > 0 \qquad \text{for all } t_i > 0, i = a, b,$$

or

$$< 0 \qquad \text{for all } t_i > 0, i = a, b.$$

An example of an independent nonadditive parallel model satisfying the conditions of Proposition 5 will be found in the section Applying Factorial Reaction Time Methods.

The most interesting ways of producing mean RT additivity at the present time are through violations of one or more of the conditions in Proposition 5. Townsend and Ashby (1983, Chap. 12, p. 374) give an additive example of an independent parallel model which violates only (ii), that is, direct nonselective influence occurs. If condition (i) on independence is abandoned one may inquire whether a nonselective influence is direct or indirect or possibly both. There is sometimes an ambiguity with regard to the directness of influence in dependent models, just as is in the serial case.

We now present two parallel dependent models which yield additivity, the first of which corresponds to a now well-known alternative to a standard serial model. It is, however, instructive to see how it relates to the above propositions. Both models obey structure (iii), the monotonicity axiom. Considerations as to how violations of (iii) might lead to additivity are discussed in Townsend and Ashby (1983, p. 376).

The Capacity Reallocation System: Positive Dependence

Consider the classic display search paradigm where a single target character is shown to an observer followed by a multicharacter array which either contains a single replication of the target or no target at all. The observer responds with a "yes" button press if the target is present and a "no" button press otherwise. Accuracy is practically perfect and mean RT is plotted for differing numbers of characters in the second array (e.g., Atkinson, Holmgren, and Juola, 1969; Townsend and Roos, 1973; a memory search paradigm where the single character and multicharacter array are reversed in temporal position as in Sternberg, 1966, would do just as well for the present discussion). This type of design may be viewed as an instance of the additive factor method where it is presumed that exactly one process comes into play when an additional character is inserted into the display. In this sense, it becomes a special case reducing to the Donderian method of subtraction (because taking out a character eliminates a process).

We shall consider processing to be "exhaustive," that is, all display items complete comparison with the target. Processing must be exhaustive on trials when the target is absent but the evidence is mixed on "yes" trials when logically it is possible for comparisons to cease when the target is discovered (e.g., Sternberg, 1975; Townsend and Roos, 1973; Townsend and Ashby, 1983, Chaps. 7, 8; Theios, Smith, Haviland, Traupmann, and Moy, 1973). Ordinarily, the usual description of $E(\mathrm{RT}; x_a, x_b)$ in terms of two different levels of x_a and x_b is not apparent in the standard way of presenting the results. However, we may fashion a description in the terms employed here by making the assumption that each serial position indexed (according to spatial location—usually left to right in a linear array) corresponds to a distinct subsystem. Admittedly, this assumption seems more plausible in the context of parallel systems than serial.

Let $x_a = 0$ or 1, representing the absence or presence of a character input to process and associated with position S_a, and similarly for $x_b = 0$ or 1. It will be convenient to return to a description of processing that was very useful in earlier papers (e.g., Townsend, 1974): the analysis of intercompletion times. It may be recalled that an intercompletion time is an actual processing time in serial systems but the difference between successive actual processing times in parallel systems. As such, we will use the term "stage" to refer to the time from one intercompletion to the next. Thus, the potential for alteration of processing rates across stages (i.e., from one intercompletion time to the next) permits stochastic dependencies to emerge in certain classes of parallel systems.

Suppose the two processes are parallel and exponential within any single stage; then the system may be characterized by the exponential rates $v_{al}(x_a, x_b)$, where l = stage of processing (i.e., l = stage 1 if no characters are yet processed and l = stage 2 if exactly one character has been processed) and as noted $x_j = 0$ or 1 depending on the absence or presence of a character in position S_j, $j = a, b$. Note that the rates are undefined if no character is in the corresponding position and that stage 2 only occurs when $x_a = 1$ and $x_b = 1$. From techniques developed elsewhere (e.g., Townsend, 1974) we may derive the mean RT for the various factor level combinations as

$$E(\mathrm{RT}; 0, 0) = t_0$$

$$E(\mathrm{RT}; 1, 0) = t_0 + \frac{1}{v_{a1}(1, 0)}$$

$$E(\mathrm{RT}; 0, 1) = t_0 + \frac{1}{v_{b1}(0, 1)}$$

$$\begin{aligned} E(\mathrm{RT}; 1, 1) &= t_0 + \frac{v_{a1}(1, 1)}{v_{a1}(1, 1) + v_{b1}(1, 1)} \\ &\quad \times \left(\frac{1}{v_{a1}(1, 1) + v_{b1}(1, 1)} + \frac{1}{v_{b2}(1, 1)}\right) \\ &\quad + \frac{v_{b1}(1, 1)}{v_{a1}(1, 1) + v_{b1}(1, 1)} \\ &\quad \times \left(\frac{1}{v_{a1}(1, 1) + v_{b1}(1, 1)} + \frac{1}{v_{a2}(1, 1)}\right). \\ &= t_0 + \frac{1}{v_{a1}(1, 1) + v_{b1}(1, 1)} \\ &\quad + \frac{v_{a1}(1, 1)}{v_{a1}(1, 1) + v_{b1}(1, 1)} \frac{1}{v_{b2}(1, 1)} \\ &\quad + \frac{v_{b1}(1, 1)}{v_{a1}(1, 1) + v_{b1}(1, 1)} \frac{1}{v_{a2}(1, 1)}. \end{aligned} \tag{9}$$

In the *capacity reallocation system*, it is assumed that when an item is completed (here, a comparison) then the capacity that had been allocated to that item is freed to be placed on the unfinished items so that in the context of a set of parallel exponential channels, the overall production rate (the summed rates on the unfinished items) remains constant until the last item is completed.

The capacity reallocation model may be derived from Eq. (9) by setting

$$v_{a1}(1, 0) = v_{b1}(0, 1) = v \qquad \text{the constant overall rate}$$

$$v_{a1}(1, 1) + v_{b1}(1, 1) = v$$

and

$$v_{a2}(1, 1) = v_{b2}(1, 1) = v.$$

It then follows that this model is equivalent to the standard serial exponential system. Note that it need not be that $v_{a1}(1, 1) = v_{b1}(1, 1)$ so that serial position effects (different RTs on different positions) are possible.

We have shown previously that this model predicts positive correlations in completion times so that independence does not hold (Townsend, 1974, p. 155; Townsend and Ashby, 1983, Chap. 4). Thus, when $n = 2$, it is proven that if S_b has completed processing, the chances are improved that S_a is also done. In any case, Eq. (9) becomes

$$E(\text{RT}; 0, 0) = t_0$$

$$E(\text{RT}; 1, 0) = t_0 + 1/v$$

$$E(\text{RT}; 0, 1) = t_0 + 1/v$$

$$E(\text{RT}; 1, 1) = t_0 + 2/v,$$

obviously satisfying additivity. Observe that during stage 1 and $n = 2$ there is a direct nonselective influence because $v_{a1}(1, 1)$ has changed (to $v - v_{b1}(1, 1)$) from what it was as $v_{a1}(1, 0)$ (i.e., v). That is, factor x_b is affecting S_a in a direct manner. On the other hand, under the reallocation concept, there is also an indirect influence in that $v_{a1}(1, 1)$ speeds up from $v - v_{b1}(1, 1)$ to $v_{a2}(1, 1) = v$. It is assumed that in parallel processing none of the rates of unfinished items are ever zero.

A Negatively Dependent Parallel System which Predicts Mean RT Additivity

We may take the same display search paradigm described in the previous section and use Eq. (9) to create the present system. Let $v_{a1}(1, 0) = v = v_{b1}(0, 1)$ as before but now let $v_{a1}(1, 1) = v_{b1}(1, 1) = v$ and $v_{a2}(1, 1) = v_{b2}(1, 1) = 2v/3$.

Plugging these values in Eq. (9) proves additivity. It is known that a decrease in rates from stage 1 ($v_{a1}(1, 1)$ and $v_{b1}(1, 1)$) to stage 2 ($v_{a2}(1, 1)$ and $v_{b2}(1, 1)$) results in a negative dependency; that is, knowing that, say S_a is done decreases the probability that S_b is finished processing (e.g., Townsend, 1974). Thus, either type of dependency can be associated with an indirect nonselective influence which induces mean RT additivity.

Collating the findings of the preceding results on serial and parallel systems vis a vis mean RT additivity, we are in a position to state the following principle. In order to express the principle in as clear a fashion as possible, the extra proviso of parallel systems, monotonicity (condition (iii) of Proposition 5) is assumed as a postulate.

PROPOSITION 6. Mean RT Additivity Parallel–Serial Duality Principle.

Serial systems. *In serial systems, the combination of selective influence and stochastic independence of processing times produces mean RT additivity whereas* *nonselective influences and/or stochastic dependence can (but need not) produce interactions.*

Parallel systems. *In parallel systems, the combination of selective influence and stochastic independence of processing times produces interactions, whereas nonselective influences and/or stochastic dependence can (but need not) produce mean RT additivity.*

We thus see that parallel and serial systems tend to respond in opposite ways when confronted with the same conditions on their structure and function, with regard to mean RT additivity, at any rate. A summary of the results for the processing systems described is given in Table 1.

The table is divided into serial (upper half) and parallel (lower half) systems. Within each of these divisions, all combinations of independence vs. dependence and selective vs. nonselective influence are represented. The fourth line indicates the possible forms of nonselective influence. Below these descriptors are the outcomes of factorial experiments which involve that type of system.

The capitalized results in Table 1 indicate the only logically possible consequence for that system. The proposition(s) relevant to this conclusion is (are) given under each result. (Because Proposition 6 represents a general principle applicable to the entire table it is not listed.) Thus, as given by Proposition 1, a selectively influenced, dependent serial system always yields additive results under factorial manipulation. By "selectively influenced" we have ruled out indirect nonselective as well as direct nonselective influence. Although logically possible, a dependent serial system that is additive is unlikely to be found in nature due to indirect nonselective influence. Similarly, a parallel independent system with selective influence is necessarily nonadditive by Proposition 5.

TABLE 1

Results of Factorial RT Experiments for Given Processing Systems

	Independent		Dependent	
	Selective Influence	Nonselective Influence	Selective Influence	Nonselective Influence
		Direct		Dir Indir Both
Serial	ADDITIVE (Propositions 1, 2)	nonadditive but could be additive (Proposition 3)	ADDITIVE (Proposition 1)	nonadditive but could be additive (Proposition 3)
Parallel	NONADDITIVE (Proposition 5)	nonadditive but could be additive	nonadditive but could be additive	nonadditive but could be additive

The remaining table entries are not so absolutely defined but do show a certain propensity for nonadditivity. The systems given as examples previously demonstrate that nonadditivity is not, however, the only logical result. In the future, we may be able to specify what additional dimensions play a role in specifying the exact nature of processing for these systems. The end result of this strategy will be the identification of a broad range of processing systems, including the parallel and serial systems developed here, via the results of factorial experiments.

Now that we know how two systems with diametrically opposed styles of processing behave with respect to the additive factor method, we focus our attention on a discussion of areas in which the method has been applied and on errors in use of the method.

Applying Factorial Reaction Time Methods

Brief Critique of the Additive Factor Literature

Appropriate applications of the additive factor method to a task or paradigm require an a priori model or system which identifies functional processes for which factors which will influence such processes can be postulated. The experimenter-theorists's procedure is then to propose a model and search (as well as hope) for factors which will influence the processes in such a way as to identify them and to validate the model. Alternatively, one may analyze the pattern of interactions or lack thereof and attempt to narrow the class of potential theoretical candidates.

Outside of standard cognitive implementations, as were mentioned in the introduction, the method is also found in such distinct settings as developmental, clinical, physiological, and social-personality research under the auspices of the information-processing paradigm. Clinical and developmental applications generally attempt to localize processing deficits or developmental effects in a particular subprocess, implicitly accepting a system much like Smith's (1968). Simon and Pouraghabagher (1978) concluded that the primary effect of aging was on the encoding stage in a choice reaction time task which manipulated stimulus discriminability and age. In one of many attempts to localize the information-processing deficit in schizophrenics, Neufeld (1977) used a sentence verification task and found that the deficit was not in the comparison process. Hemsley (1976) compared the performance of schizophrenics and depressives on a card-sorting task (with mean times ranging from 31 to 98 s) and concluded that schizophrenics showed difficulties in response selection.

Psychophysiologists often reference RT models in conjunction with other dependent measures. Gomer, Spicuzza, and O'Donnell (1976) recorded both RTs and P300 latencies and found evidence that the P300 latency is indicative of the time required for initial sensory processing. Lacey and Lacey (1980), however, associate the P300 waveform with decisional processes. Van der Molen and Orlebeke (1980) compared the effect of auditory signal intensity on heart rate and choice reaction time, concluding that both output processes and stimulus encoding processes may be influenced by the arousal effect of loud auditory signals.

Rogers (1974) proposed a separate 3-process model of responding to personality items—stimulus encoding, stimulus comprehension, and binary decision. See Fig. 5. Evidence for additivity of item length, item ambiguity, and item controversiality were obtained.

While the logic applied often appears to be a model test via factor manipulations, the actual practice more closely resembles use of information regarding the additivity and interactions of factors to construct a model of serial processes and generate relevant labels for those processes. In practice, this generally produces the relatively unenlightening finding of additivity of processes which on logical grounds alone must be distinct and sequentially ordered and probably widely separated physiologically as well. For instance, stimulus intensity and response discriminability fall into this category. It would seem our goal should be results that carry more information. Along similar lines, use of multiple factor experiments is often encouraged since they may provide more conclusive discrimination between possible process arrangements.

Taylor (1976) went further in holding that the more appropriate strategy is not to search for additive factors, but rather for interactions. The justifications for such an approach are, first, the basic inadequacy of statistical tests in attempts to accept the null hypothesis (see later sections for more specific remarks) and, also, the possibility of inappropriate conclusions if, for example, factors interact (i.e., influence a common process) in an additive manner. He proposes examination of patterns of interactions across a series of experiments in order to more clearly delineate the processes involved. We note, though, that processes within Taylor's scheme continue to be sequential, although not strictly serial (1976, p. 176). The distinction between these two is that "sequential" implies a directed information flow which does not exclude the possibility of overlapping or concurrent processes, while a "strictly serial" system would not allow two or more processes to be in execution at the same time.

The advantage of such a strategy is the ability to obtain more information for producing a functional definition of the processes involved, due to the larger base of factor effects considered. A survey of the literature indicated a dearth of studies developing and exploring a "pattern of interactions" within the same investigation, much less across experiments. Also, when more than two or three factors are manipulated, the results, while interesting, may be difficult to interpret (e.g., Yio, 1978).

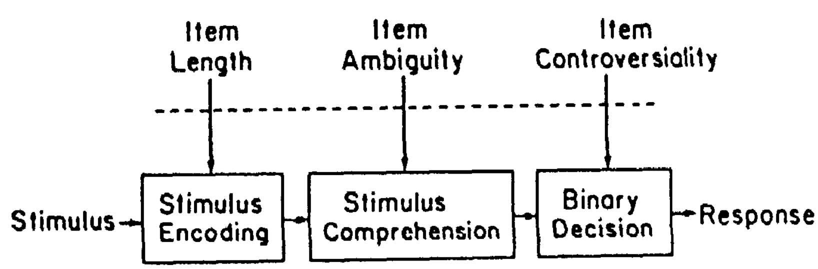

FIG. 5. The model of responding to personality items proposed by Rogers (1974).

We look to future specification of systems via knowledge of the subadditivity or superadditivity of interactions as well as possibly more complex forms of interaction. Further, ordering of processes, that is, designating which processes precede others, which to date has been assumed by convention or intuition, seems more definitive and verifiable via certain of Schweickert's methods (1978; 1983).

The diversity of applications has led to an even greater diversity in factors which appear as independent variables in such experiments. Aside from the standard factors such as stimulus quality and number of response alternatives, less directly controllable factors such as age, mental capacity, and diagnosis of schizophrenia are also used. Sternberg's original definition of a factor was "an experimentally manipulated variable, or a set of two or more related treatments called "levels"; the effect of a factor is the change in the response measure induced by a change in the level of that factor." (1969a, pp. 282–283) One of the potential characteristics of a process (i.e., "stage" in the earlier terminology) also constrained its output to be unaffected by factors influencing its duration. Factors which influence not only durations but outputs of processes have not been investigated and would certainly complicate (if not eradicate) any additive factors interpretation. (See Townsend and Ashby, 1983, Chap. 12, pp. 401–412 for some discussion of this issue.) Whether any variable which can be construed as having more than one level is appropriately viewed as a factor is open to question and certainly requires a more precise definition of factor effects for a given process in a mathematical sense.

The dependent variable originally proposed for the methodology was RT, although transmitted information (Kantowitz & Knight, 1974) and P300 latency (McCarthy & Donchin, 1981) have been employed. The validity of these and other measures in terms of the methodology have yet to be systematically explored. Minimal requirements appear to be that such measures can be regarded as random variables and are on at least an interval scale, due to the use of the arithmetic mean statistic.

In fact, mathematically tractable use of additive factors logic requires the use of mean RTs, rather than, say, medians. The expectation (i.e., the mean) of a sum of random variables (in this case serial processing times) is equal to the sum of the expectations, which is true for any set of random variables whether or not they are independent. However, the same is not true for the medians, unless they are equal to the means, which usually occurs if the density function is symmetric. This is not a reasonable assumption due to the skewed nature of reaction time distributions (e.g., Ratcliff and Murdock, 1976). Still, analyses using medians and other transformations which destroy additivity appear (e.g., Silverman, 1973; Acosta and Simon, 1976; and Fairweather and Hutt, 1978).

Pachella (1974) provides a critical analysis of the additive factor method, specifically cautioning against interpretations of RTs which do not take error data into account, due to the possibility of speed-accuracy trade-offs. RTs are generally considered the minimum time required to complete certain mental processes, but if speed requirements are emphasized (either by the experimenter of the subject's internal criterion), RTs may precede completion of processing. This is generally presumed to lead to more errors for shorter RTs and evidence of a speed-accuracy trade-off. The relevant point is thus that it is not proper to draw conclusions regarding processing models from RTs which may not involve complete processing.

Heeding this advice, reports of error data have become more common, although clearly not universal. Even so, there is still some ambiguity in the definition of errors. Errors carry many interpretations: premature responses, incorrect responses in discrimination tasks, inappropriate tripping of voice relays, various RT outliers, or misidentification of targets in priming studies. Typically, average error rates are under 5% and most experimenters will thus exclude error RTs from the data.

The issue of statistically testing for speed-accuracy trade-offs was not covered by Pachella (1974), and many experimenters have likewise disregarded it. For others, it is sufficient that decreasing RTs do not raise error rates, or that they are "low" and basically similar. A high positive correlation of RTs and percent error were taken to be evidence against any latency-error trade-off by Clark & Chase (1974) and Howard (1975), although this may be thought debatable at best. An analysis of variance (ANOVA) is often performed, although there seems to be a deficit in interpretations of what would or would not imply a speed-accuracy trade-off. Are nonsignificant main effects for the error rates sufficient (as in Shulman & McConkie, 1973) or must the interaction also be nonsignificant (as in Barron, 1975)?

One may also consider what interpretation should be made of error rates which increase with RT, which is common enough (e.g., Holender & Bertelson, 1975) and could possibly give a significant result under statistical testing. Note that evidence against a speed-accuracy trade-off could arise from either constant error rates or error rates which increase with RTs.

With regard to mode of analysis and presentation, a perusal of the literature shows a preponderance of RT graphs, some tables of RTs, fewer error rate reports, and only an occasional report of standard errors or standard deviations—all components of a convincing argument for the processing system proposed. For an even more definitive interpretation, Sternberg (1969a) advised analysis of higher cumulants.[6] Such analyses, which provide information regarding the stochastic independence of processes are very rare in the literature (as exceptions, see Biederman & Kaplan, 1970; and Blackman, 1975).

[6] Cumulants can be written in terms of moments about the mean. The first cumulant is the first moment about the mean, which is zero. Some higher cumulants κ_i, in terms of moments about the mean μ_i, are

$$\kappa_2 = \mu_2$$
$$\kappa_3 = \mu_3$$
$$\kappa_4 = \mu_4 - 3\mu_2^2.$$

Note that the second moment about the mean μ_2 is equal to the variance. In general, cumulants $\kappa_1, \kappa_2, \ldots, \kappa_r$ are defined by the identity in t

$$\exp\left\{\kappa_1 t + \kappa_2 \frac{t^2}{2!} + \cdots + \kappa_r \frac{t^r}{r!} + \cdots\right\} = 1 + \mu_1 t + \mu_2 \frac{t^2}{2!} + \cdots + \mu_r \frac{t^r}{r!} + \cdots,$$

where μ_r is the rth moment about an arbitrary point a. (Kendall & Stuart, 1963). For discussion of the use of cumulants and related statistics, see Ratcliff (1979) and Ashby and Townsend (1980).

Statistical Questions

For current purposes, the most relevant (and disturbing) point with regard to use of additive factors logic centers on the statistical tests involved and their interpretations. Recently, Rouanet, Levine, & Holender (1978) have proposed the use of Bayes-fiducial methods for validating an additive model of binary choice RT while a nonparametric method using RT distributions was proposed by Ashby & Townsend (1980). Still, ANOVA is clearly the most widely (verging on universally) used technique, although Sternberg (1969a) also proposed the use of interaction contrasts. Such applications and the resulting interpretations certainly demand much closer scrutiny than they have thus far received.

In most cases, ANOVA is used for factorial experiments of at least the second order, most often a repeated-measures design with subjects as a random factor. Typically, the basic unit for a particular cell of the ANOVA table is a mean RT for an individual subject over trials in a session, sometimes averaged over sessions as well. The number of trials contributing to a mean ranges from 5 to 600 (typically about 160) and the number of sessions is generally under 5. The number of subjects in such an experiment is usually around 15, with a range of 5 to 40.

Appropriate use of the ANOVA for the methodology requires obtaining significant main effects for the factors. Generally, a significance level of 0.01 or 0.05 is used for such a test. The object of interest is then the test for interaction of the factors, which most often is similarly tested at an 0.01 significance level.

The null hypothesis for the interaction in the ANOVA is that $\sigma^2_{ab} = 0$ for factors x_a and x_b. Classical hypothesis testing procedures seek to reject the null hypothesis at an alpha (α) level which guarantees a small Type I error, i.e., a probability α that the null hypothesis is wrongly rejected. The problem is that additive factors implementations typically have not sought to reject the null but, rather, to accept it, since acceptance has been taken to imply distinct serial processes, while rejection has, in the past, implied virtually nothing. In particular, in using a small α to lower Type I error, the experimenter concurrently raises the probability of falsely accepting the null hypothesis (Type II error). It has been proposed that a more acceptable significance level for the interaction test would be 0.25 or 0.50 to increase the power of the test, but this caveat has certainly not been heeded in the literature.

The importance of maximizing the power of tests for interactions in factorial experiments clearly has been ignored. Whether this disregard has led to models which have powerless tests as their basis is certainly worth investigating and will be examined for one particular experiment in the following section.

Testing a Nonadditive Parallel Model Against Data and Considerations of Statistical Power

In an attempt to begin investigation of the difficulties associated with the additive factor method and analysis of variance and, particularly, to determine the ability of the method to distinguish between classes of additive serial and "interactive" parallel systems, a parallel model meeting the conditions of one of the impossibility theorems (Proposition 5) was constructed for testing under the standard procedures.[7]

Even when, as it seems in the example below from the literature, the supposed psychological processes should logically be sequential (not necessarily strictly serial), it is of value to test whether a diametrically opposed type of system (i.e., an independent parallel system, for example) can make predictions statistically similar to the data, or for that matter to serial additive systems. If a parallel categorically nonadditive system cannot be *empirically* distinguished from a serial additive system by the data, we may be in trouble. Thus, the *general* class of nonoverlapping sequential (or other) systems will probably permit predictions even closer to those of strict serial systems.

This section has four goals. The first is to construct a parallel model that cannot predict mean RT additivity. The second is to see how well such a model can come to fitting statistics from a representative experiment from the literature, in fact an experiment where additivity was accepted. We will then subject the predictions of the best fitting parallel model to a standard ANOVA type of mean RT additivity analysis. The fourth goal is to examine the way power of the ANOVA changes as a function of sample size for two different α levels.

A model of two independent parallel processes in series with a residual process was constructed. (See Fig. 6.) The RT under consideration was of course the exhaustive completion time (i.e., time to finish both) of the processes S_a and S_b, plus the residual processing time. The residual component was given a mean of 160 ms and a standard deviation of 36 ms, which seems to be an adequate estimate of simple RT (Woodworth & Schlosberg, 1954).

In an attempt to treat a more general type of distribution than had thus far been used, each of the parallel processes was given a gamma distribution, written

$$g(t) = \frac{v}{k!}(vt)^{k-1} e^{-vt}, \qquad t > 0,$$
$$= 0, \qquad t \leqslant 0.$$

The parameter v is the rate of processing and k is the number of subprocesses within the process. Note that when $k = 1$ the distribution is exponential.

Proposition 5 specifies conditions to be met to render mean RT additivity impossible. The first, stochastic independence of processing times, is met since

$$\begin{aligned} G_{ab}(\tau_a, \tau_b; x_a, x_b) &= P(T_a \leqslant t_a, T_b \leqslant t_b; x_a, x_b) \\ &= P(T_a \leqslant t_a; x_a, x_b)\, P(T_b \leqslant t_b; x_a, x_b) \\ &= G_a(\tau_a; x_a, x_b)\, G_b(\tau_b; x_a, x_b) \end{aligned}$$

[7] The term "interactive parallel" is meant to imply that nonadditive (interactive) effects are produced in mean RTs, not that two parallel processes interact with one another in real time. In fact, as was seen earlier, independent parallel processes produce interactions under broad conditions; whereas parallel systems which are stochastically dependent can produce mean RT additivity.

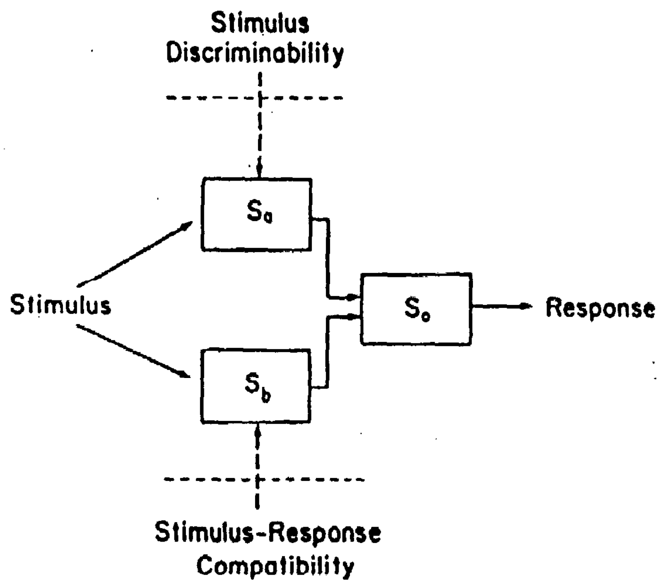

FIG. 6. The independent parallel system proposed as an alternative to a serial additive system of the data of Biederman & Kaplan (1970).

Selective influence holds by construction, since we assume the factor effects are manifested in the rate parameter. The distributions for processes S_a and S_b have their own rate parameters v_a and v_b, respectively, i.e.,

$$G_a(\tau_a; x_a, x_b) = G_a(\tau_a; x_a) = G_a(\tau_a; v_a)$$
$$G_b(\tau_b; x_a, x_b) = G_b(\tau_b; x_b) = G_b(\tau_b; v_b).$$

The third condition requires monotonicity of factor effects in the distribution functions. If we examine the relationship of the distributions given above and the factors, it can be shown that $G_a(\tau_a; v_a)$ is a monotonic increasing function of v_a for any fixed τ_a and k_a and the same holds for v_b and $G_b(\tau_b; v_b)$. The respective means are monotonic decreasing functions of v_a and v_b.

As an added condition, we note that part (B) of Proposition 5 is also satisfied, which states that if the monotonicity is in the same direction for both, subadditivity occurs. Computational implementation of the model, described below, exemplified this assertion in that every set of predictions of the model in a 2×2 factorial design produced a negative interaction, i.e., a subadditive result. (See Fig. 7 for one such graph.)

The computational formulas for the model, given in Appendix A, yielded a mean and variance for a given system, without the residual process, for each particular 2×2 factor combination. The residual processing time mean and variance were then added to these computed values to get the overall system mean and variance. In order to provide a linkage with experimental work, the mean will represent the expected value of the sample mean RT for n homogeneous subjects each in m trials. The variance is similarly the average variance of n homogeneous subjects in m trials.

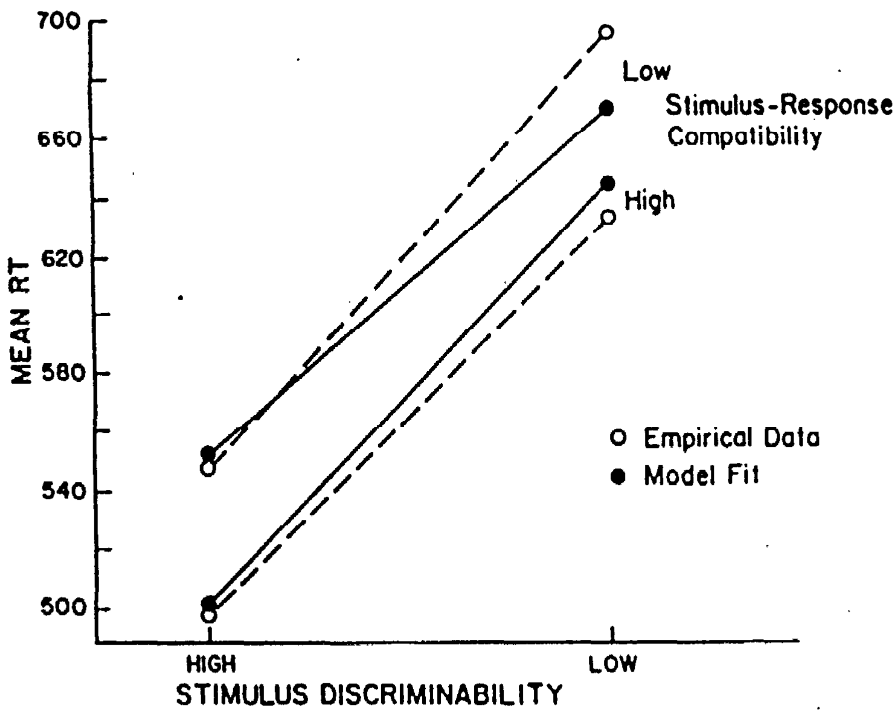

FIG. 7. The best chi-square fit of the parallel system to the empirical data of Biederman & Kaplan (1970).

Since we average the variance of homogeneous subjects, this is more easily seen as the variance of one subject over m trials.

The model of this system was fit to the data of Biederman and Kaplan (1970), one of the most competent applications of the method we located and which included reports and analysis of higher cumulants as well as means. They used a brightness discrimination task in a 2×2 design which showed additivity of stimulus discriminability and stimulus—response (S–R) compatibility. There were 16 subjects and 48 trials per subject per condition with two sessions.

It might be objected that if we submit this parallel model to ANOVA, we are violating the usual ANOVA assumptions. The appropriate rebuttal is that if an underlying empirical system is in fact parallel (or something else) then the data from an experiment on that system does in fact violate the assumptions in exactly the same way; therefore, this is the very tack we should wish to take.

The fit to both means and standard deviations concurrently was obtained via a STEPIT (Chandler, 1965) routine which minimized the chi-square value. The rate parameters were permitted to vary within the routine and were assumed to be functions of the factor levels. The number of subprocesses k_a and k_b were treated as free parameters but did not vary across factor levels. It is interesting that while varying the number of subprocesses produced a barely observable effect on the means-only fit, decent fits to both means and standard deviations required manipulation of the number of subprocesses (i.e., variation of k_a and k_b). There were therefore 6 parameters in the model, with only the rates reflecting factor levels vs.

eight data points, the four means and four standard deviations. Note that even if there were more parameters in the model than there were data points, the independent parallel structure of the model (and other conditions of Proposition 5) would prevent a perfect fit to truly additive data.

The best fit found had $k_a = 4$ and $k_b = 20$ with a chi-square value of 2.00 with two degrees of freedom. See Table 2 for the relevant numerical results and Fig. 7 for the predicted and observed mean RTs. The error in fit was statistically insignificant at the 0.05 level.

There are several worthwhile points with regard to the model test and fit. First, if one only compared the best model predictions with the mean RT data in the graph of Fig. 7, one would have concluded erroneously that the fit was completely unacceptable; yet statistically the model fit quite well according to the conventional criteria, even demanding predictions for variance as well as means. Further, the parameters are within a reasonable range of variation. Note too, that we have in a way biased the situation against the parallel models because we chose data that in fact evidence some degree of superadditivity, (as the reader can see in Fig. 7) while the model must

TABLE 2

Biederman & Kaplan (1970) Empirical Data (in ms)

C = S–R Compatibility
D = Stimulus Discriminability

	Means		Standard Deviations	
	HiD	LoD	HiD	LoD
LoC	548.00	696.00	94.87	184.95
HiC	500.00	634.00	84.81	179.63

Model Predictions:
Two independent, gamma-distributed processes plus a residual component.

	HiD	LoD	HiD	LoD
LoC	553.93	670.46	95.16	178.30
HiC	501.85	646.84	84.49	189.01

Parameters of Best Fit of Model:

Number of Subprocesses
for process S_a, $k_a = 4$
for process S_b, $k_b = 20$

Processing rates
$v_a(1) = 0.009$ = rate of process S_a under LoD
$v_a(2) = 0.274$ = rate of process S_a under HiD
$v_b(1) = 0.051$ = rate of process S_b under LoC
$v_b(2) = 0.059$ = rate of process S_b under HiC

Model Fit:

$df = 2$, $\alpha = 0.05$
Best $\chi^2 = 2.00$ vs. $\chi^2(0.05) = 6.00$

predict subadditivity. On the other hand, an investigator might be willing to use the data superadditivity as prima facie evidence against independent parallel processing, but then it must also be taken as qualitative evidence against pure serial additivity as well.

We now turn to ANOVA analyses of the Biederman and Kaplan (1970) data and the data predictions of our best-fitting parallel model. Can this model, which provided a good statistical fit to the data, be rejected by a traditional ANOVA test of additivity? Does it, with its inherent subadditivity, give as close an approximation to mean RT additivity as does the superadditive data, relative, of course, to the pertinent variances?

Note that a statistical model is not to be confused with a model of a processing system, such as given in Fig. 6. Each component of a statistical model identifies a source of variance in the gathered data, such that an ANOVA can be performed to determine which sources are significant. The additive factor method extends the statistical results so as to use the existence of significant factor effects to implicate particular processing arrangements.

A plausible statistical model for the ANOVA of Biederman and Kaplan's (1970) data would include subjects and sessions as components, as well as all the interactions of these with the other factors; however, the relatively simple processing system espoused was based on an assumption of homogeneity of subjects and a disregard for any practice effects, so such components were not included. In this respect, the ANOVA to be performed on our system's predictions will not exactly duplicate the ANOVA done by Biederman and Kaplan. However, we feel it is justifiable and instructive to examine identically run ANOVAs for both our predictions and the actual data insofar as the statistical signficance of the factor effects, which is identifiable in both, is our primary focus.

As such, the statistical model for the ANOVA was

$$x_{ijkl} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{lijk}$$

$$\begin{aligned} i &= \text{factor level for factor } x_a && (1, 2) \\ j &= \text{factor level for factor } x_b && (1, 2) \\ k &= \text{subject} && (1, n) \\ l &= \text{trial} && (1, m), \end{aligned}$$

where x_{ijkl} is a single RT trial for each subject under a specific experimental condition. The general procedure for performing an ANOVA under such conditions is to use subject means as the cell entries, which implies the error variance is the variance between subjects within a cell. Because that variance is zero, under our assumptions, the variance for a subject within a cell was used as the error variance.

The procedure used for calculation of this 2×2 ANOVA is given in Appendix B. The information required was available from the model predictions. The number of trials per subject was 96, i.e., 48 trials per session times 2 sessions. Table 3 gives the

TABLE 3

Biederman & Kaplan (1970) Discriminability × Compatibility

	MS	*F*	
ANOVA on Empirical Data			
D	1908576.00	92.350	***
C	290400.00	14.051	***
D × *C*	4704.00	0.227	
ERROR	20666.62		
ANOVA on Parallel Model Predictions			
D	1641425.05	78.434	***
C	137531.76	6.571	**
D × *C*	19439.32	0.928	*
ERROR	20927.41		

Notes. * = significant at $\alpha = 0.50$, ** = significant at $\alpha = 0.05$, *** = significant at $\alpha = 0.01$. $F_{50}(1, 380) = 0.456$, $F_{95}(1, 380) = 3.866$, $F_{99}(1,380) = 6.702$.

results of the ANOVA for the independent parallel model, as well as the same procedure applied to the empirical data for comparative purposes.

The results for the empirical data, under both modified and simplified calculations, verify the Biederman & Kaplan (1970) results for the mean, i.e., the main effects are significant but the interaction is not. In fact, we note that the main effects were significant at an $\alpha = 0.01$ while the interaction was *not* significant at $\alpha = 0.01$ through $\alpha = 0.50$, even stronger evidence for additivity than supplied by Biederman & Kaplan.

The model's results were different. Again, main effects were significant at $\alpha = 0.01$, but, while the interaction was not significant at $\alpha = 0.01$ or 0.05, it *was* significant at $\alpha = 0.50$. The impact of these results is then, that out best-fit independent parallel model was indeed distinguishable from "additive" data via statistical testing, but *not* at the signficance levels typically used in the literature.

From an analogous point of view, it would be worthwhile in the future to consider the relationship to power and relevant α level of *serially* arranged, non-normally distributed times (i.e., time distributions ordinarily employed by RT theorists or those described by experimenters which are decidedly non-normal.) For instance, under what circumstances might such a model yield nonadditivity in a statistical sense?

The statistical power of the *F* test of the interaction was determined for the ANOVA of Biederman and Kaplan's actual data as given in Table 3. Because of our interest in the distinguishability of additive data and parallel models, the best-fit independent parallel model was used as the alternative hypothesis for the power analysis.

Cohen (1969) provides tables and procedures for determining the power of an

TABLE 4

Power of the F Test*

	α	
m	0.05	0.10
96	0.183	0.286
200	0.326	0.445
300	0.446	0.557
400	0.539	0.636

Notes. *These values are linear interpolations of Cohen's tables which he termed an adequate approximation to intermediate values; degrees of freedom = 1, effect size = 0.073, m = number of trials, α = significance level.

interaction for an ANOVA test.[8] Specification of the degrees of freedom for the interaction, the number of entries per cell, the significance level, and an effect size are required. The degrees of freedom for an interaction are $(p-1)(q-1)$, where p and q represent the number of rows and columns, respectively, and are one for Biederman & Kaplan's 2 × 2 factorial experiment. The number of entries per cell is $M = 96$ (i.e., trials per condition), as given in the ANOVA. The significance level used is $\alpha = 0.05$. The effect size is interpreted as a measure of variability of (interaction) effects, whose mean is zero, standardized by the common within (cell) population standard deviation (Cohen, 1969, p. 365). It is obtained via the following formula in terms of alternate-hypothetical cell means

$$f = \frac{\sigma_y}{\sigma},$$

where σ_y is the variability of interaction effects and σ is the within-cell population standard deviation. The particulars of this calculation are given in Appendix C and resulted in an effect size (f) equal to 0.073, which is interpreted as small.

The power of this test, or equivalently one minus the probability of Type II error, is 0.183. (See Table 4.) In other words, this means that the probability of falsely accepting additivity (since the null hypothesis is that $\sigma_{ab} = 0$) is 0.817, a rather wide margin of error, to say the least. Table 4 shows that either increasing the significance level (α) or increasing the number of trials (m), or both, will increase the power of the test. In particular, note that 300 trials are required to achieve a power of more than 0.50 at an α level of 0.10.

All three of the previous analyses, the chi-square model fit, the ANOVA tests, and the power analysis point to the same conclusions. Under the typical model testing and statistical procedures, systems that are diametrically opposed to those which

[8] Koele (1982) gives a general overview of power calculations for analysis of variance under Models I, II, and III.

most naturally predict additivity can produce behavior in the range accepted as additive. Thus, the usual techniques and criteria might not reject additive (e.g., serial) models on the basis of data coming from a *parallel* nonadditive system. Conversely, it appears that a model of such a parallel system can be acceptably close to data taken as supportive of serially arranged subsystems. An alternative approach might be to use confidence intervals. This treatment possesses certain advantages over the traditional test of the null hypothesis as pointed out, for example, by Myers (1966).

On the bright side is the fact that with larger α levels or greater sample sizes, the parallel model was rejected and, more critically, was rejected at a smaller α level than was the "additive" data. This points up again the importance of testing alternative explanations or models against one another, especially considering the usual uncertainty regarding the "proper" level of significance and sample size (see also Snodgrass and Townsend, 1980, on this issue).

Of course, an underlying parallel system might behave in a way that was more obviously nonadditive relative to the variances. To enhance this possibility, more extreme factor levels should be used when feasible; this would, for instance, make subadditivity more salient.

Summary and Conclusion

The analysis of RT additivity and interactions via factorial manipulation is a potentially powerful tool to help elucidate covert mental processes. However, paucity of knowledge of the predictions of viable systems other than the classic independent serial system and severe statistical weakness in typical applications have impeded true scientific advance gained with these techniques.

In particular, although we have recently begun to learn more about the factorial predictions of some alternatives to standard serial processing (e.g., Taylor, 1976; McClelland, 1979; Ashby and Townsend, 1980; Schweickert, 1978), information regarding pure parallel processing with regard to factorial issues has been lacking. We have proved analytically (i.e., in closed mathematical form and without the aid of computer simulation) that there exists a large and natural class of parallel models which cannot predict perfect mean RT additivity. One might well claim that if we are unable to discriminate pure nonadditive parallel processing from true serial additive processing then we are in trouble virtually everywhere.

Indeed, we found that an exemplary nonadditive parallel model gave a good fit to data formerly accepted as additive. Further, an ANOVA treatment of the parallel generated "data" found that it was additive at the traditional criteria. A power analysis confirmed that only with enhanced power due to a larger α and/or larger sample size could the account by the parallel system be rejected.

The implications are clear. More mathematically rigorous knowledge is required concerning psychologically interesting processing systems, as well as better comprehension of the power of our statistical methods relative to alternative process explanations. It is likely that most conclusions from the literature are theoretically weak because of low power and because of the a priori high likelihood of the sequentiality (not necessarily pure seriality) of the processes chosen for study. Finally, more thought should be given in applications of the factorial technique to testing alternative models against one another, rather than simply accepting or rejecting the null hypothesis as affirming or falsifying a single type of system.

APPENDIX A

Consider two processes S_a and S_b which each have a gamma-distributed completion time with, respectively, k_a and k_b component subprocesses and processing rates v_a and v_b. The mean reaction time for a parallel exhaustive model of two such processes in which the order of intercompletion of the subprocesses is not deterministic can be formulated in terms of two possible cases. (See Fig. 8.)

The reaction time for such an exhaustive system is the time required until both processes are completed, the maximum of the two processing times in parallel systems. Case I is that in which process S_a finishes prior to process S_b, which means all k_a subprocesses are completed and some number of the k_b subprocesses have yet to be completed. Case II shows process S_b finishing prior to process S_a and when all k_b subprocesses are completed, some of the k_a subprocesses are not. There are clearly many possible orderings of the intercompletions, each with their own probability of occurrence.

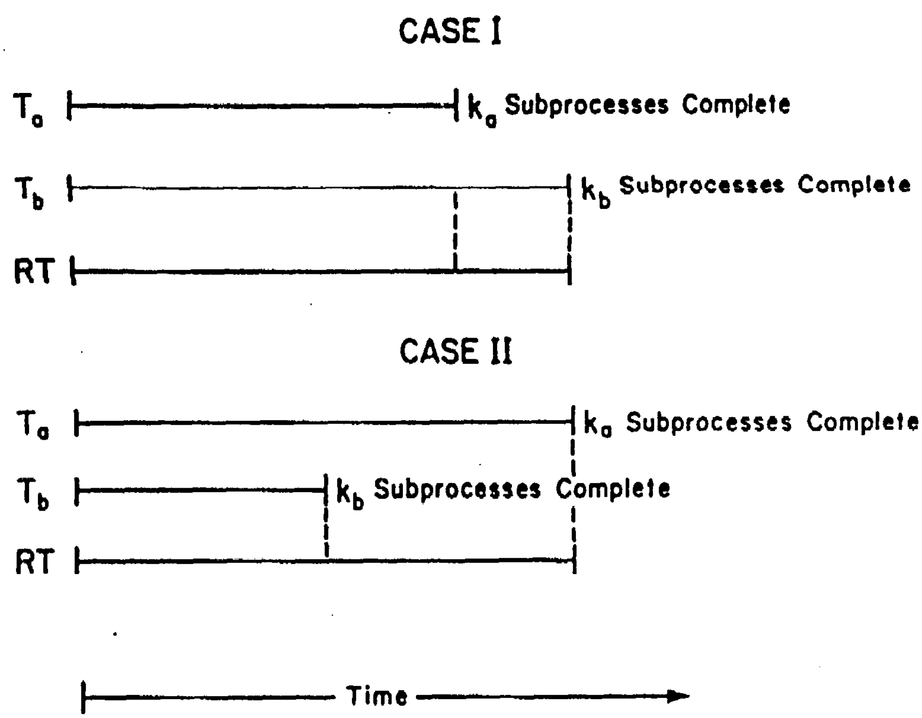

FIG. 8. The two cases of intercompletion orderings of two parallel processes used for the development of the independent, gamma-distributed parallel system, given in Appendix A.

We characterize the formula for mean reaction time as

$$E(T) = \sum_{\substack{\text{possible}\\ i \text{ stages}}} P\begin{pmatrix}S_a \text{ finishes}\\ \text{before } S_b\\ \text{and } i \text{ sub-}\\ \text{processes}\\ \text{of } S_b \text{ are}\\ \text{completed}\end{pmatrix}\left[\begin{pmatrix}\text{Time to complete}\\ k_a \text{ and } i \text{ sub-}\\ \text{processes with}\\ \text{both } S_a \text{ and } S_b\\ \text{processing}\end{pmatrix} + \begin{pmatrix}\text{Time to complete}\\ \text{the remaining}\\ \text{subprocesses}\\ \text{with only } S_b\\ \text{processing}\end{pmatrix}\right]$$

$$+ \sum_{\substack{\text{possible}\\ j \text{ stages}}} P\begin{pmatrix}S_b \text{ finishes}\\ \text{before } S_a\\ \text{and } j \text{ sub-}\\ \text{processes}\\ \text{of } S_a \text{ are}\\ \text{completed}\end{pmatrix}\left[\begin{pmatrix}\text{Time to complete}\\ k_b \text{ and } j \text{ sub-}\\ \text{processes with}\\ \text{both } S_a \text{ and } S_b\\ \text{processing}\end{pmatrix} + \begin{pmatrix}\text{Time to complete}\\ \text{the remaining}\\ \text{subprocesses}\\ \text{with only } S_a\\ \text{processing}\end{pmatrix}\right].$$

The probability that S_a finishes before S_b and i subprocesses of S_b are processed in that time can be represented as a negative binomial probability for i "failures" (i.e., intercompletions for S_b) before k_a "successes," where the probability of "success" is $v_a/(v_a + v_b)$ and "failure" is $v_b/(v_a + v_b)$. Similarly for the case in which S_b finishes before S_a.

We, therefore, write the formula for mean reaction time

$$E(T) = \sum_{i=0}^{k_b-1} \binom{k_a+i-1}{i}\left(\frac{v_a}{v_a+v_b}\right)^{k_a}\left(\frac{v_b}{v_a+v_b}\right)^{i}\left(\frac{k_a+i}{v_a+v_b}+\frac{k_b-i}{v_b}\right)$$
$$+ \sum_{j=0}^{k_a-1} \binom{k_b+j-1}{j}\left(\frac{v_b}{v_a+v_b}\right)^{k_b}\left(\frac{v_a}{v_a+v_b}\right)^{j}\left(\frac{k_b+j}{v_a+v_b}+\frac{k_a-j}{v_a}\right).$$

The formula for the variance is

$$\mathrm{Var}(T) = E(T^2) - [E(T)]^2,$$

where

$$E(T^2) = \frac{(k_b+1)(k_b)}{(v_b)^2}\left(\frac{v_a}{v_a+v_b}\right)^{k_a}\sum_{i=0}^{k_b+1}\frac{(k_a+i-1)!}{(k_a-1)!(i)!}\left(\frac{v_b}{v_a+v_b}\right)^{i}$$
$$+\frac{(k_a+1)(k_a)}{(v_a)^2}\left(\frac{v_b}{v_a+v_b}\right)^{k_b}\sum_{j=0}^{k_a+1}\frac{(k_b+j-1)!}{(k_b-1)!\,(j)!}\left(\frac{v_a}{v_a+v_b}\right)^{j}.$$

A more general development of computational formulas for the mean and variance of various arrangements of processes which are distributed as the sum of exponential random variables is given in Fisher and Goldstein (1983).

As an illustration of the mean calculations, suppose that process S_a has $k_a = 2$ subprocesses and process S_b has $k_b = 1$. The rates are v_a and v_b, respectively. Then,

$$E(T) = \sum_{i=0}^{0}\binom{2+i-1}{i}\left(\frac{v_a}{v_a+v_b}\right)^{2}\left(\frac{v_b}{v_a+v_b}\right)^{i}\left(\frac{2+i}{v_a+v_b}+\frac{1-i}{v_b}\right)$$
$$+\sum_{j=0}^{1}\binom{1+j-1}{j}\left(\frac{v_b}{v_a+v_b}\right)^{1}\left(\frac{v_a}{v_a+b_b}\right)^{j}\left(\frac{1+j}{v_a+v_b}+\frac{2-j}{v_a}\right)$$
$$=\binom{1}{0}\left(\frac{v_a}{v_a+v_b}\right)^{2}\left(\frac{v_b}{v_a+v_b}\right)^{0}\left[\frac{2}{v_a+v_b}+\frac{1}{v_b}\right] \quad \text{(i)}$$
$$+\binom{0}{0}\left(\frac{v_b}{v_a+v_b}\right)^{1}\left(\frac{v_a}{v_a+v_b}\right)^{0}\left[\frac{1}{v_a+v_b}+\frac{2}{v_a}\right] \quad \text{(ii)}$$
$$+\binom{1}{1}\left(\frac{v_b}{v_a+v_b}\right)^{1}\left(\frac{v_a}{v_a+v_b}\right)^{1}\left[\frac{2}{v_a+v_b}+\frac{1}{v_a}\right] \quad \text{(iii)}. \qquad (11)$$

If we label the subprocesses for S_a as a_1, a_2 and for S_b as b_1, we note that term (i) represents finishing order $\langle a_1, a_2, b_1\rangle$; (ii) is $\langle b_1, a_1, a_2\rangle$; and (iii) is $\langle a_1, b_1, a_2\rangle$, an exhaustive listing of the possible finishing orders.

As further illustration of the result, consider (11)(iii) and note that associated with the order $\langle a_1, b_1, a_2\rangle$ the expected value may be represented, in terms of *intercompletion* times of subprocesses (rather than completion times).

$$E[T_{a1}+T_{b1}+T_{a2}\cap\langle a_1, b_1, a_2\rangle]$$
$$=\int_{t_{a1}=0}^{\infty}\int_{t_{b1}=0}^{\infty}\int_{t_{a2}=0}^{\infty}(t_{a1}+t_{b1}+t_{a2})\,g(t_{a1}, t_{b1}, t_{a2})\,dt_{a1}\,dt_{b1}\,dt_{a2},$$

wherein

$$g(t_{a1}, t_{b1}, t_{a2}) = (v_a e^{-v_a t_{a1}} e^{-v_b t_{a1}})\cdot(v_b e^{-v_b t_{b1}} e^{-v_a t_{b1}})\cdot(v_a e^{-v_a t_{a2}}).$$

Thus,

$$E[T_{a1} + T_{b1} + T_{a2} \cap \langle a_1, b_1, a_2 \rangle]$$

$$= \iiint t_{a1} v_a e^{-(v_a+v_b)t_{a1}} v_b e^{-(v_a+v_b)t_{b1}} v_a e^{-v_a t_{a2}} \, dt_{a1} \, dt_{b1} \, dt_{a2}$$

$$+ \iiint t_{b1} v_a e^{-(v_a+v_b)t_{a1}} v_b e^{-(v_a+v_b)t_{b1}} v_a e^{-v_a t_{a2}} \, dt_{a1} \, dt_{b1} \, dt_{a2}$$

$$+ \iiint t_{a2} v_a e^{-(v_a+v_b)t_{a1}} v_b e^{-(v_a+v_b)t_{b1}} v_a e^{-v_a t_{a2}} \, dt_{a1} \, dt_{b1} \, dt_{a2}$$

which, after tedious computations, yields

$$E[T] = \frac{v_a v_b}{(v_a + v_b)^3} + \frac{v_a v_b}{(v_a + v_b)^3} + \frac{v_b}{(v_a + v_b)^2} = \frac{v_a v_b}{(v_a + v_b)^2} \left[\frac{2}{v_a + v_b} + \frac{1}{v_a} \right]$$

which is equal to (11)(iii). Similar results are obtainable for parts (11)(i) and (11)(ii). $E(T^2)$ is derived as (this time in terms of *completion* times)

$$E(T^2) = \int_{t_a=0}^{\infty} \int_{t_b=t_a}^{\infty} t_b^2 \, g_{ab}(t_a, t_b) \, dt_b \, dt_a + \int_{t_b=0}^{\infty} \int_{t_a=t_b}^{\infty} t_a^2 \, g_{ab}(t_a, t_b) \, dt_a \, dt_b,$$

where

$$g_{ab}(t_a, t_b) = g_a(t_a) \, g_b(t_b) = \frac{v_a}{(k_a - 1)!} (v_a t_a)^{k_a - 1} e^{-v_a t_a} \frac{v_b}{(k_b - 1)!} (v_b t_b)^{k_b - 1} e^{-v_b t_b}.$$

Integration yields the given result.

Appendix B

Assume an equal number m of entries in each cell and a 2×2 factorial experiment, i.e., $p = 2$ and $q = 2$. For ease of notation, the two factors x_a and x_b will be labeled C and D, i.e., S–R compatibility and stimulus discriminability, and are fixed. The following data are available:

		D 1	D 2
C	1	$\bar{X}_{11\ldots}$, VAR$_{11\ldots}$,	$\bar{X}_{12\ldots}$, VAR$_{12\ldots}$,
	2	$\bar{X}_{21\ldots}$, VAR$_{21\ldots}$,	$\bar{X}_{22\ldots}$, VAR$_{22\ldots}$,

The row means, $\bar{X}_{1\ldots}$, and $\bar{X}_{2\ldots}$; column means, $\bar{X}_{.1..}$, and $\bar{X}_{.2..}$; and grand mean, $\bar{X}_{\ldots}$, can be computed from the above data.

For a 2 × 2 factorial experiment the mean squares are defined as

$$\mathrm{MS}_c = 2m \sum_{i=1}^{2} (\bar{X}_{i\ldots} - \bar{X}_{\ldots\ldots})^2$$

$$\mathrm{MS}_d = 2m \sum_{j=1}^{2} (\bar{X}_{.j..} - \bar{X}_{\ldots\ldots})^2$$

$$\mathrm{MS}_{cd} = m \sum_{i=1}^{2} \sum_{j=1}^{2} (\bar{X}_{ij..} - \bar{X}_{i\ldots} - \bar{X}_{.j..} + \bar{X}_{\ldots\ldots})^2$$

$$\mathrm{MS}_{err} = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{m} \frac{(\bar{X}_{ij\cdot 1} - \bar{X}_{ij..})^2}{4(m-1)}$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(\text{within-cell variance for a subject})}{4}$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\mathrm{VAR}_{ij..}}{4}.$$

Note that subjects are ignored due to the assumption of homogeneity. Also, the MS_{err} is the average of the variances for each cell as fit by the model. The F ratios are

$$F_c = \frac{\mathrm{MS}_c}{\mathrm{MS}_{err}} = \frac{8m[(\bar{X}_{1\ldots} - \bar{X}_{\ldots\ldots})^2 + (\bar{X}_{2\ldots} - \bar{X}_{\ldots\ldots})^2]}{\mathrm{VAR}_{11..} + \mathrm{VAR}_{12..} + \mathrm{VAR}_{21..} + \mathrm{VAR}_{22..}}$$

$$df = (1, 4m - 4)$$

$$F_d = \frac{\mathrm{MS}_d}{\mathrm{MS}_{err}} = \frac{8m[(\bar{X}_{.1..} - \bar{X}_{\ldots\ldots})^2 + (\bar{X}_{.2..} - \bar{X}_{\ldots\ldots})^2]}{\mathrm{VAR}_{11..} + \mathrm{VAR}_{12..} + \mathrm{VAR}_{21..} + \mathrm{VAR}_{22..}}$$

$$df = (1, 4m - 4)$$

$$F_{cd} = \frac{\mathrm{MS}_{cd}}{\mathrm{MS}_{err}} = \frac{4M[(\bar{X}_{11..} - \bar{X}_{1\ldots} - \bar{X}_{.1..} + \bar{X}_{\ldots\ldots})^2 + (\bar{X}_{12..} - \bar{X}_{1\ldots} - \bar{X}_{.2..} + \bar{X}_{\ldots\ldots})^2 + (\bar{X}_{21..} - \bar{X}_{2\ldots} - \bar{X}_{.1..} + \bar{X}_{\ldots\ldots})^2 + (\bar{X}_{22..} - \bar{X}_{2\ldots} - \bar{X}_{.2..} + \bar{X}_{\ldots\ldots})^2]}{\mathrm{VAR}_{11..} + \mathrm{VAR}_{12..} + \mathrm{VAR}_{21..} + \mathrm{VAR}_{22..}}$$

$$df = (1, 4m - 4).$$

Appendix C

To find the effect size,

$$f = \frac{\sigma_y}{\sigma},$$

wherein

$$\sigma = \sqrt{\sum^{pq} y_{ij}^2/[(p-1)(q-1)+1]}$$

and

$$y_{ij} = \bar{X}_{ij\cdot\cdot} - \bar{X}_{i\cdot\cdot} - \bar{X}_{\cdot j\cdot\cdot} + \bar{X}_{\cdot\cdot\cdot\cdot}$$

where the $\bar{X}$s are from the alternate-hypothetical cell means, main effect means, and grand mean.

The within-cell population standard deviation was approximated by the averaged standard deviation over all cells, i.e.,

$$\sigma = \frac{\sum^p \sum^q \sqrt{\mathrm{VAR}_{ij}}}{4}.$$

Computations gave

$$f = \frac{10.06}{136.74} = 0.073.$$

References

Acosta, E., Jr. & Simon, J. R. (1976). The effect of irrelevant information on the stages of processing. *Journal of Motor Behavior*, 8, 181–187.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Ashby, F. G. (1982). Testing the assumptions of exponential, additive reaction time models, *Memory & 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*, 21, 93–123.

Ashby, F. G. & Townsend, J. T. (Submitted for publication). *Varieties of perceptual independence.*

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

Barron, R. W. (1975). Locus of the effect of a distinguishing feature in a memory search task. *Memory & Cognition*, 3, 302–310.

Biederman, I., & Kaplan, R. (1970). Stimulus discriminability and S–R compatibility: Evidence for independent effects in choice reaction time. *Journal of Experimental Psychology*, 86, 434–439.

Blackman, A. R. (1975). Test of the additive-factor method of choice reaction time analysis. *Perceptual and Motor Skills*, 41, 607–613.

Chandler, J. P. (1965). Subroutine Stepit: An algorithm that finds the values of the parameters which minimize a given continuous function. A copyrighted program: J. P. Chandler. Quantum Chemistry Program Exchange. Bloomingon, IN: Indiana University.

CLARK, H. H., & CHASE, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory & Cognition*, 2, 101–111.

COHEN, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

DONDERS, F. C. (1969). On the speed of mental processes. Translated by W. G. Koster. In W. G. Koster (Ed.), *Attention and performance II*. Amsterdam: North-Holland.

FAIRWEATHER, H., & HUTT, S. J. (1978). On the rate of gain of information in children. *Journal of Experimental Child Psychology*, 26, 216–229.

FISHER, D. L., & GOLDSTEIN, W. M. (1983). Reaction time models of complex cognitive tasks: PERT networks, service completion diagrams and statistics of the reaction time distribution. *Journal of Mathematical Psychology*, 27, 121–151.

GOMER, F. E., SPICUZZA, R. J., & O'DONNELL, R. D. (1976). Evoked potential correlates of visual item recognition during memory-scanning tasks. *Physiological Psychology*, 4, 61–65.

HEMSLEY, D. R. (1976). Attention and information-processing in schizophrenia. *British Journal of Social and Clinical Psychology*, 15, 199–209.

HOLENDER, D., & BERTELSON, P. (1975). Selective preparation and time uncertainty. *Acta Psychologica*, 39, 193–204.

HOWARD, J. H., JR. (1975). The attentional demands of negation in a memory-scanning task. *Memory & Cognition*, 3, 319–324.

KANTOWITZ, B., & KNIGHT, J. L., JR. (1974). Testing, tapping, time-sharing. *Journal of Experimental Psychology*, 103, 331–336.

KENDALL, M. G., & STUART, A. (1963). *The advanced theory of statistics* (2nd ed.). London: Griffin.

KOELE, P. (1982). Calculating power in analysis of variance. *Psychological Bulletin*, 92, 513–516.

KUHN, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

LACEY, B. C., & LACEY, J. I. (1980). Cognitive modulation of time dependent primary bradycardia. *Psychophysiology*, 17, 209–221.

MCCLELLAND, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.

MCCARTHY, G., & DONCHIN, E. (1981). A metric for though: A comparison of P300 latency and reaction time. *Science*, 211, 77–79.

MYERS, J. L. (1966). Fundamentals of Experimental Design. Boston: Allyn and Bacon.

NEUFELD, R. W. J. (1977). Components of processing deficit among paranoid and nonparanoid schizophrenics. *Journal of Abnormal Psychology*, 86, 60–64.

PACHELLA, R. G. (1974). The interpretation of reaction time in information-processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.

PARZEN, E. (1960). *Modern probability theory and its applications*. New York: Wiley.

PIETERS, J. P. M. (1983). Sternberg's additive factor method and underlying psychological processes: Some theoretical considerations. *Psychological Bulletin*, 93, 411–426.

RATCLIFF, R. ((1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86, 446–461.

RATCLIFF, R., & MURDOCK, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83, 190–214.

ROGERS, T. B. (1974). An analysis of the stages underlying the process of responding to personality items. *Acta Psychologica*, 38, 205–213.

ROUANET, H., LEVINE, D., & HOLENDER, D. (1978). Model acceptability and the use of Bayes-fiducial methods for validating models. In J. Requin (Ed.), *Attention and performance VII* (pp. 687–701). Hillsdale, NJ: Erlbaum.

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

SCHWEICKERT, R. (1983). Latent network theory: Scheduling of processes in sentence verification and the Stroop effect. *Journal of Experimental Psychology: Human Learning & Memory*, 9, 353–383.

Bulletin, 69, 77–110.

SNODGRASS, J. G., & TOWNSEND, J. T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 330–354.

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

STERNBERG, S. (1969a). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276–315.

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

STERNBERG, S. (1975). Memory scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology*, 27, 1–32.

TAYLOR, D. A. (1976). Stage analysis of reaction time. *Psychological Bulletin*, 83, 161–191.

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: Advances in research and theory: VII*. New York: Academic Press.

THEIOS, J., SMITH, P. G., HAVILAND, S. E., TRAUPMANN, J., & MOY, M. C. (1973). Memory scanning as a serial self-terminating process. *Journal of Experimental Psychology*, 97, 323–336.

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. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Erlbaum.

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

TOWNSEND, J. T. (1976b). A stochastic theory of matching processes. *Journal of Mathematical Psychology*, 14, 11–52.

TOWNSEND, J. T. (1981a). *In pursuit of nonadditive reaction time theories*. Paper presented at 14th Annual Mathematical Psychology Meeting, Santa Barbara, August.

TOWNSEND, J. T. & PIOTROWSKI, S. M. (1981b). *Parallel systems and factorial additivity of mean reaction time*. Paper presented at Psychonomics Society Meeting, Philadelphia, November.

TOWNSEND, J. T. (1981c). Some characteristics of visual whole report behavior. *Acta Psychologica*, 47, 149–173.

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

TOWNSEND, J. T., & ROOS, R. N. (1973). Search reaction time for single targets in multiletter stimuli with brief visual displays. *Memory & Cognition*, 1, 319–332.

VAN DER MOLEN & ORLEBEKE, J. F. (1980). Phasic heart rate change and the *U*-shaped relationship between choice reaction time and auditory signal intensity. *Psychophysiology*, 17, 471–481.

WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental psychology*. NY: Holt, Rinehart, & Winston.

YIO, J. H. (1978). *Visual Recognition of Words vs. Nonwords*. Unpublished doctoral dissertation, Purdue University.

RECEIVED: June 15, 1984

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium

[redacted]