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

Vol. 39, No. 4, December 1995
Printed in Belgium

Spatio-temporal Properties of Elementary Perception: An Investigation of Parallel, Serial, and Coactive Theories*

JAMES T. TOWNSEND
Indiana University

AND

GEORGIE NOZAWA
Dartmouth College

A mathematical theory and related experimental methodology are developed that permit strong, converging tests of parallel versus serial versus channel summation processing and of exhaustive versus self-terminating processing. An experimental design is used for two studies, in which the presence or absence of a target in each of two positions, is factorially combined with two levels of brightness (a version of the double factorial paradigm). When both targets are present (redundant target condition) the two levels of brightness permit factorial technology to determine mental architecture and stopping rules. Comparison of the single versus double target conditions allows capacity analyses that strongly reinforce the double target factorial phase of the investigation. The results provide decisive support for parallel channels with either a self-terminating stopping rule or a coactive summation of information. General serial models and any variety of exhaustive processing were conclusively falsified.

We previously developed a taxonomy for elementary cognitive processes (e.g., Townsend, 1974; Townsend & Ashby, 1983) based on the notions of architecture (e.g., serial vs parallel), capacity (e.g., unlimited vs limited vs super capacity), stopping rules (e.g., self-terminating: can stop when sufficient information in a given task is accrued vs exhaustive: always processes all incoming information), and stochastic dependence vs independence. The present theory relies on these concepts but requires the incorporation of a new one, "coactivation vs separated processing," as suggested by Miller (1982, 1988). The reason is that the principle of coactivation is based on the consolidation of activation into fewer channels and the decision is based on activation there, for instance the reaching of preset criteria, rather than on the original separate processors. In our simple situation, the coactivation carries activation from two parallel channels into a single channel, and processing stops when a criterion is reached. Note that the original channels possessed a parallel architecture and the earlier taxonomy did not encompass coactivation properties.

Prototypical schemata for serial, parallel, and coactive processing are shown in Figs. 1a, 1b, and 1c, respectively. We view A and B as being the items that are processed on the current trial by, respectively, processes or channels X and Y. For simplicity, we associate the experimental factors, X and Y, as discussed below, with their associated processors as shown in the figure.

Observe that in Figs. 1a and 1b processing is completed on each separate channel or item, whereas in Fig. 1c the parallel activations from A (in X) and B (in Y) are summed, or otherwise consolidated in general, before a detect decision is made. The bypass in Fig. 1a around the Y process simply allows for self-termination after the X process is completed. Of course, in an exhaustive serial process, that branch will be absent or inoperable. Similarly, in true parallel processing the logical function might operate as an "and" gate (exhaustive processing) or an "or" gate (self-termination), but in neither serial nor parallel processing does all the information or activation converge to a common channel before a decision is reached. Notice that the concept of stopping rule does not apply to coactive processing when all information is consolidated into a single channel.

One central aspect of the present investigation is the further extension of what we have called systems factorial technology (Townsend, 1992). This technology is oriented around psychological systems identification by way of associating experimental factors with distinct psychological processors (subsystems, operations, etc.). Thus, Fig. 1 indicates a pair of experimental factors X and Y selectively influencing the X and Y processors. Systems factorial technology includes the Dondersian (1868) subtraction

Correspondence concerning this article should be addressed to James T. Townsend, Indiana University, Department of Psychology, Bloomington, IN 47405. E-mail: jtownsen@indiana.edu.

* The authors are grateful for the very helpful comments of F. Gregory Ashby, Asher Cohen, Hans Colonius, Ehtibar Dzhafarov, Howard Egeth, Howard Hughes, Wolfgang Schwarz, and Richard Schweickert on a previous version of this paper. Valuable critiques were also contributed by Jeff Miller, Lester Krueger and two anonymous reviewers. We are extremely appreciative of the care and industry put forth by Sheryl Mobley in preparing the manuscript. The research has been supported by NSF Grant #8710163 to the first author.

0022-2496/95 $12.00

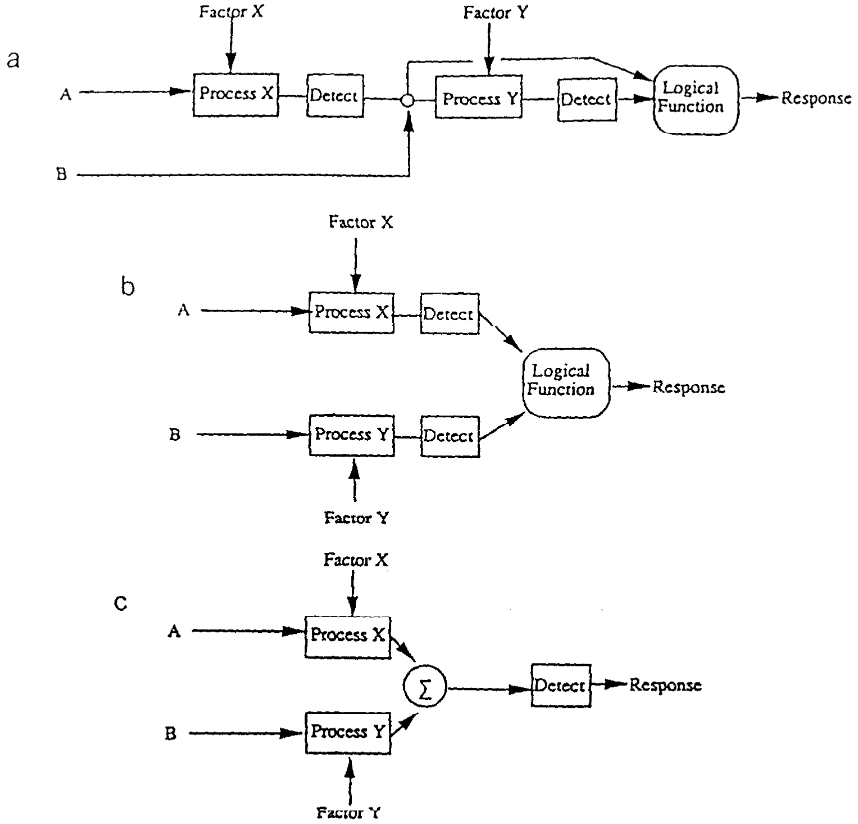

FIG. 1. Schematics of (a) a serially arranged system, (b) a parallelly arranged system, and (c) a channel summation coactivation model. All X and Y subsystems are selectively influenced by their respective experimental factors.

method, Sternberg's (1969) additive-factor method, Schweickert's latent network theory (1978, 1982, 1985; Schweickert & Wang, 1993), the trichotomy method (Schweickert & Townsend, 1989; Townsend & Schweickert, 1985; Townsend & Schweickert, 1989), and the factorial branch of Townsend and colleagues' stochastic modeling theory (e.g., Townsend, 1984; Townsend & Ashby, 1983, Chapter 12). The major branch of this technology resides within reaction time (hereafter RT) methodology as applied to elementary cognitive processes (e.g., Egeth, 1966; Eriksen & Spencer, 1969; Estes & Taylor, 1964; Garner, 1974; Hick, 1952; Krueger, 1978; Luce, 1986; Neisser, 1967; Posner, 1978; Schweickert, 1992; Smith, 1968; Sternberg, 1966; Theios, 1973; Townsend & Ashby, 1983; Welford, 1980). A number of terms pertinent to this and other topics are qualitatively defined in Table 1.

The present experimental application is to the visual detection of one or two dots and attempts to determine the architecture, whether, if the architecture is parallel, processing is ordinary parallel or coactive, and the stopping rule (if applicable, as in serial or parallel processing). Also, capacity both is assessed in its own right and serves to help in the other identifications. We build on earlier work by the authors and their colleagues (Nozawa, 1989; Townsend, 1981; Townsend & Nozawa, 1988; Townsend & Piotrowski, 1981). The theory and implied methodology developed herein have begun to be applied in other venues (Hughes, Reuter-Lorenz, Nozawa, & Fendrich, 1994; Nozawa, Reuter-Lorenz, & Hughes, 1994).

One facet of the double factorial paradigm is the well-known "redundant target design." Consider a task where one or more signals or patterns are designated as targets and the observer is to make a certain response if one or more of these appears in a display and a different response when no target is presented. The experimenter is interested in the effects on RT of trials where a single target is presented versus trials where more than one target is presented (e.g., Bernstein, 1970; Colonius, 1990a; Egeth & Mordkoff, 1991; Grice, Canham, & Boroughs, 1984; Nickerson, 1973; Raab, 1961; Snodgrass & Townsend, 1980). In fact, a speeding

TABLE 1

Theoretical Glossary

Additive-factor method	The branch of systems factorial technology in which emphasis is on additivity versus any kind of interaction. If factors are additive (that is, indicate zero interaction) then the conclusion is the determination of separate serially arranged processes. If an interaction is found it is supposed that both factors influence the same subprocess or, factor selective influence fails.
Capacity	A theoretical construct pertaining to influences on performance when the load on a system is varied. For example, the number of items to be processed constitutes a change in load. Even if all the items are targets, capacity can be measured by parameters or functions within a processing system.
Channel summation model	A special variety of coactivation model where information in each parallel channel comes in the form of discrete counts. It is supposed that these counts are combined by summation in the output.
Coactive models	A parallel architecture which assumes that input from the separate parallel channels is consolidated into a resultant common processor.
Double factorial paradigm	An experimental paradigm wherein two levels or types of stimuli in two positions are combined with factorial manipulation of another variable such as intensity.
Exhaustive processing	The system processes all items in an array. For example, all channels available are completed even though some of the information or items may be redundant.
Factorial selectivity	A postulate in most branches of systems factorial technology is that separate factors influence separate processes within a processing system.
Grice's inequality	Under conditions stated in the text, the survivor function for the redundant target condition is bounded above by a certain function of the survivor functions from the single target conditions. Violations of this bound imply very limited capacity within the present theory.
Interactive race model	A parallel model in which non-stochastically independent channels may communicate in a sensory or decisional phase.
Superposition (neural summation) model	A special coactivation model that is also a special case of channel summation, where the counts in the parallel channels are produced by Poisson stochastic processes.
Mean interaction contrast	The interaction or contrast produced by two or more factors at the level of the mean statistic.
Miller's inequality	Under certain conditions, the survivor function for the redundant target condition is bounded below by a combination of survivor functions from the two single target conditions. Violation of this bound implies supercapacity within the present theory.
Probability summation (statistical advantage)	The statistical advantage of two or more signals rather than a single signal when the processing times are arranged in parallel. A special case occurs when the processing channels are stochastically independent.
Redundant target paradigm	When multiple targets are presented to a observer and any one of the individual targets is sufficient to determine a correct response. The targets are said to be redundant.
Self-termination	The system ceases processing as soon as a single target is located. A special case of self-termination is minimum time processing within a redundant target paradigm.
Stopping rules	Rules to determine when a network stops processing. Special cases of interest are self-termination and exhaustive processing.
Survivor interaction contrast	The interaction contrast equation is applied to survivor functions (one minus the cumulative distribution function) rather than the mean RT. Thus an interaction contrast *function* rather than a single number is produced.
Systems factorial technology	A growing methodology that emphasizes identification of process structure through manipulation of experimental factors, typically under the assumption of factorial selectivity.

up on trials with *multiple* (i.e., redundant) targets typically occurs, relative to trials with fewer targets. We will discuss this facet more later. A related literature studies the effects that identical, similar or dissimilar surrounding distractors have on RT or accuracy (e.g., Bjork & Murray, 1977; Estes, 1972; Shapiro & Krueger, 1983). In those investigations, unlike in the present experiments, the additional symbols typically are supposed to be ignored by the observer.

The double factorial paradigm, in a prototypical visual realization with two potential targets in each of two positions, presents a target in neither position, a single target in one of the positions or in the other, or a target in both positions. The participant gives an affirmative response if the target appears in either position and a negation only if the target appears in neither position. Thus, it can be seen that the target presence or absence in the two positions is both a factorial manipulation and a redundant target design. The second factor of the double factorial paradigm appears in an experimental variation of a second variable such as intensity or salience that affects RT. These two factors are crossed with each other; hence the name of the design.

Theorems that specify differential predictions for the various types of models are derived in the Appendices and discussed qualitatively in the text. These theorems and the tests they provide are new. Another novel development is that of predictions on the entire cumulative distribution functions of RT. The cumulative distribution function, $F(t)$, or distribution function in brief, is the integral of the probability density function $f(t)$. The survivor function, $S(t)$, defined as $S(t) = 1 - F(t)$, where F is the cumulative distribution function, is a major variable of interest in the investigation. Although the distributional results imply predictions at the level of mean RTs, they are much stronger than the latter.

In addition, capacity of the mental operations under study is assayed with the help of several recently proposed RT inequalities and our present derivation of a capacity measure based on hazard functions. Predictions with regard to this capacity measure for parallel and coactivation models and the relations of these to the major inequalities are derived. The theory and its implications are then combined with other current mathematical work to provide bounds on the processing capacity evidenced by our observers.

It cannot be overemphasized that the factorial tests, which take place within data from trials where both targets are present, are logically independent of the capacity analyses relating to single vs double target trials. This double-pronged attack permits much stronger convergence of evidence than is possible with either alone.

BRIEF HISTORICAL PERSPECTIVE

It may be useful for nonspecialists to briefly review the history. Others may wish to skip to the next section. In 1868, Donders published a groundbreaking paper on the "method of subtraction." This method postulated that by running different experimental conditions with-versus-without a particular cognitive operation and then subtracting RT for the simpler task from that for the more complex, an estimate of the average time consumed by the more complex component could be obtained. This implicitly assumes strict seriality of the psychological components and, of course, the ability to delete them experimentally. Although based on very strong assumptions, the method still forms a useful part of the experimental battery (e.g., Ashby & Townsend, 1980; Gottsdanker & Schrag, 1985; Kohfeld, 1969; Meyer & Schvaneveldt, 1971).

A century later, S. Sternberg revitalized the general approach with his "additive-factor method" (1969). Basically, this method, like Donders', also assumed that psychological processes were laid out in a nonoverlapping temporal (perhaps, but not necessarily, physical) series of stages (see Fig. 1). However, rather than postulating that the experimenter could add or subtract entire processes out of the system, it was supposed that experimental factors might be found that could selectively affect specific processes. If this occurred then, given satisfaction of certain conditions, the mean RTs should show no interactions; that is, the factors would have additive effects. This outcome is pictured in Fig. 2 along with the other two cases, under- and overadditivity. Underadditivity is found when the additional prolongation caused, say by factor X, is less when factor Y is at the level where it already prolongs RT rather than at the "faster" level. Overadditivity is just the reverse. A result of additivity would be taken as supporting the existence of separate (and serial by definition) processes. Thus, the name "additive factors." On the other hand, if the factor effects were interactive, it was concluded that the factors did not selectively influence separate processes. This could mean either that only one process was in fact operating instead of several or that the assumption of "selective influence" was wrong.

The additive factor technique was clearly significantly more general than Donders' subtractive strategy and it has become an extremely popular method of demonstrating separate psychological processes. The reader is referred to Ashby & Townsend (1980), Pachella (1974), Pieters (1983), Schweickert (1978), Taylor (1976), Theios (1973), Townsend (1984), and Townsend & Ashby (1983) for further discussion of the additive factor method.

Perhaps the most imposing limitation of the additive factor method is its confinement to serially arranged

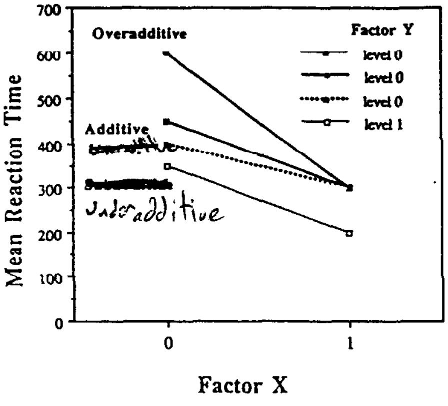

FIG. 2. Hypothetical experimental results from a 2×2 factorial experiment. The bottom line is kept constant to compare the various prototypical results of underadditivity, additivity and overadditivity shown in the three lines above it.

systems. Although Sternberg (1969) suggested that giving up seriality could lead to interactions, no rigorous theory of how various types of architectures would behave under factorial manipulation existed until recently. It turns out that underadditivity is the natural prediction in the case of exhaustive parallel processing. Suppose that RT is prolonged by manipulating one of the factors which slows up the processing time of one channel (process, item etc.). Then it can be predicted that the slowdown effect caused by the second factor will be less than that before the first factor was changed (Piotrowski, 1983; Schweickert, 1978, 1982, 1983; Townsend & Ashby, 1983; Townsend & Piotrowski, 1981). Figure 2 shows underadditivity, defined by the smaller increment in RT caused by either factor, if the other factor had already led to an increase.

One of the theoretical vanguards in factorial technology has been the expansion to quite sophisticated mental networks, much of it spearheaded by Schweickert (e.g., Schweickert, 1978, 1982, 1985; Schweickert & Townsend, 1989; Townsend & Schweickert, 1985; Townsend & Schweickert, 1989). Work by Fisher and Goldstein has been instrumental in providing for numerical explorations and model tests of such networks (e.g., Fisher & Goldstein, 1983; Goldstein & Fisher, 1991, 1992).

The redundant target part of the paradigm, involving comparisons of single versus double target presentation, does not follow from interaction analyses of previous factorial theories (e.g., Schweickert, 1978; Schweickert & Townsend, 1989; Sternberg, 1969). The former relies on the single factor manipulation of target multiplicity and examination of performance effects, whereas the latter involves the manipulation of two experimental factors, typically at the same processing load (e.g., same number of presented items).

IMPLEMENTATION OF THE DOUBLE FACTORIAL PARADIGM

Consider now the double factorial paradigm in more detail. We combine the idea of possibly different stimuli in two display positions with the factorial manipulation of a second variable, such as light intensity of the individual stimuli. For instance, each of two stimulus positions could contain Stimulus A or B. Suppose the response rule is "respond *YES* if either position contains Stimulus A, otherwise respond *NO*." Then, when both stimuli are A, if perceptual processing is self-terminating, the first position to be completed can determine the response. The latter is referred to as "minimum time" processing. On the other hand, a trial with both positions containing a B would demand that exhaustive processing of both stimuli occur. Because of the double 2×2 factorial manipulations, this constitutes the double factorial paradigm, as mentioned earlier.

We concentrate on two aspects of the double factorial paradigm: The first aspect features a brightness manipulation when both targets are present and analyzes the factorial interaction contrast. The second part compares the two types of conditions, trials with one target present and trials with both targets present.

The trials with both positions containing stimulus A constitute a form of the redundant target condition, noted above, since the stimulus A's are redundant in the sense that either is sufficient to determine a correct response. Several versions of the general redundant target paradigm have become popular in testing parallel versus serial and self-terminating versus exhaustive processing because many parallel self-terminating models predict a decrease in RT as the number of items increases. The latter prediction is exceedingly difficult for serial models and exhaustive models of either type to make (e.g., Egeth, Folk, & Mullin, 1988; Fournier & Eriksen, 1990; Meijers & Eijkman; 1977; Mordkoff & Yantis, 1991; Townsend, 1974; van der Heijden, La Heij, & Boer, 1983). Redundant target designs have also been useful in studies employing accuracy as a dependent variable (e.g., Colquhoun, 1975; Kinchla, 1974, 1977; Shaw, 1982; Townsend, 1968).

However, the double factorial paradigm itself has been utilized only a few times. It has the advantage that the redundant target aspect, in conjunction with the factorial stimulus-type aspect, can yield converging evidence on processing architecture and stopping rule, and capacity all within the same paradigm (Nozawa 1989, 1992; Townsend & Nozawa, 1988). Various forms of the double factorial paradigm can be produced by manipulating the stimulus–response assignment rules. Thus, instead of responding *YES* if either position contains an A and *NO* only if both positions contain a B (the basic form used in the present study), other possibilities arise. Townsend & Piotrowski (1981; see also Piotrowski, 1983) originally developed a version in which each of the four stimuli was assigned an unique response and presented some of the mathematical theory behind it. Egeth and colleagues have also begun exploring this method (Egeth & Dagenbach, 1991), with *X*'s and 0's of high versus low visual quality. Then, an *X* in either position required one response whereas two 0's led to the other.

Most implementations of the additive factor method seem to have maximized the chances of observing serial stages in operation (e.g., see Nozawa, 1989; Piotrowski, 1983; Townsend, 1984). In contrast, the present experiment sets up conditions where parallel processing seems to be a natural possibility. Basically, it uses one or two briefly presented dot stimuli on each trial, and if either dot is displayed, the observer responds "yes"; only if neither is displayed does the observer respond "no." Thus, the presence or absence of dots is one factorial variation. The other factor is, in the present case, the brightness of the dots which are present. The empirical side of the study may be of interest within the area of binocular versus monocular perception and these connections will be made clear under General Summary and Discussion. However, the main purpose of the experiments is to test and explore the methodology growing out of our information processing theory.

Part I of the theory develops the orthogonal factorial design when both dots are present. Manipulating the brightness of the two dots, we may calculate the statistical interaction contrast. Let the brightness designations be "l" for low and "h" for high, as applied to positions "L" for left and "R" for right. Let us define $\overline{\mathrm{RT}}(\mathrm{l}, \mathrm{l})$ to be the mean RT of the condition under which the position L is at level l and the position R is also at level l. Suppose that increasing the brightness of a given dot reduces the mean RT. We define *mean interaction contrast*, or simply mean interaction contrast, as follows:

$$\overline{\mathrm{RT}}(\mathrm{l}, \mathrm{l}) - \overline{\mathrm{RT}}(\mathrm{h}, \mathrm{l}) - \overline{\mathrm{RT}}(\mathrm{l}, \mathrm{h}) + \overline{\mathrm{RT}}(\mathrm{h}, \mathrm{h}).$$

If the above quantity is zero, we have additivity. Otherwise we have an interaction. We have generalized this concept to operate at the level of survivor function, where more power for testability is attained.

We are concerned with factor additivity plus two types of interaction: one in which the interaction is negative is termed "underadditive" and another in which the interaction is positive is termed "overadditive" (see Fig. 2). These three qualitative properties provide for testing parallel versus serial versus coactive and within parallel and serial models, self-termination versus exhaustive processing.

Part II of the theoretical section examines the capacity issue as it occurs in any redundant target experiment where the number of targets is varied across trials. The capacity issue plays a major role whenever human processing mechanisms are subjected to an increased load, even when an efficient parallel processor might be able to improve performance due to the redundant information. When two signals rather than one are presented to the visual system, potentially the individual rate of processing could decline, indicating a capacity limitation. Thus, there could be a tradeoff between limitations of increased load and improvements due to redundancy.

Suppose that the two stimuli are both targets (i.e., redundant) and the observer need only detect one of them; and it happens that the individual channel rates are unchanged. That is, capacity is not limited. Then the statistical advantage of having two signals rather than one present can lead to faster RTs (e.g., Blake, Sloan, & Fox, 1981; Colonius, 1990b; Grice, Canham, & Boroughs, 1984; Raab, 1961; Townsend & Ashby, 1983). This statistical advantage of two signals rather than one, when the processing time distributions are independent and do not change, is often called "probability summation" or "statistical advantage."

Classes of models based on separate but not necessarily statistically independent channels are generally referred to as "horse race" or simply "race" models. When faster RTs are found with redundant signals or targets, the phenomenon is sometimes called "race gains." Parallel race concepts continue to play a major role in psychophysical (e.g., Blake, Martens, Garrett, & Westendorf, 1980) and in more cognitive operations such as attention (e.g., Bundesen, 1990). Finally, it can be shown that even with mild capacity limitations, a statistical horse race advantage can occur.

Alternatively, it could be that under certain conditions, there is some form of information, or energy, consolidation (e.g., Townsend, 1968) or coactivation (Miller, 1982) that produces performance even superior to that expected on the basis of a statistical improvement. It turns out (see below) that a generic prediction of a large class of coactive models is that based on the same type of incoming channel information; they predict faster response times than parallel race models. Thus, relatively speaking, coactive models are more super capacity than unlimited capacity parallel race models. Conversely, it is possible that the increase from one to two (or more) targets leads to such large capacity decrements in individual item processing efficiency, that RTs are substantially less than what could be expected from unlimited capacity parallel race models. Such capacity decrements are possible in principle, even if the channels converge in a coactive manner.

Tests proposed by Miller (e.g., 1982) relative to super capacity (in our interpretation) and by Grice (e.g., Grice, Canham, & Gwynne, 1984) relative to very limited capacity involve inequalities at the level of distributions, or survivor functions in our interpretation. These inequalities provide bounds on RTs and by mathematical inference on capacity as defined by an index developed below. This index, or coefficient, is of interest in its own right and will provide for interesting theoretical comparisons and measures of capacity within model classes.

Some writers have taken the violation of Miller's inequality as operationally defining "coactivation" whereas satisfaction of it has been identified with "race models" (e.g., Miller, 1982). However, the foundational theoretical concepts associated with coactivation have always involved some type of combination of information (presumably gathered primitively in parallel) compared with a decision based on single channel information. On the other hand, race models imply parallelity with decisions made on separate channels, first and foremost. It is intuitively obvious though that coactive systems might actually satisfy Miller's inequality because of limited capacity in moving from one to two targets (e.g., by channel leakage) or even be worse than performance predicted by unlimited independent parallel processing. Similarly, there is no inherent reason that parallel race models could not exhibit super capacity, perhaps by interchannel communication or cross talk.

These possibilities have in fact cropped up in recent years, including the interesting interactive race model of Mordkoff & Yantis (1991), which receives more discussion below.

Our preference is therefore to separate architecture from capacity notions, but to seek to understand how various capacity effects, including super or very limited capacity, might occur with various architectures. This tack seems more realistic in avoiding overcommitment to certain theoretical predispositions (e.g., assuming that violation of Miller's inequality is equivalent to coactive combining of channel information, when this need not be true) through adherence to inelastic and oversimplified operational definitions. With this intent in mind, Colonius & Townsend (1992) propose a general mathematical classificatory scheme for parallel race models and coactivation models. The present treatment is compatible with that taxonomy.

THE THEORY AND MODELS

The two parts of the theoretical section will present the main predictions in an informal way. A brief qualitative discussion of the major assumptions underlying the propositions follows. The rigorous developments appear in the Appendices.

In the present study a serial process would scan sequentially from the input to one visual field to the other, either exhaustively or in a self-terminating manner (see Fig. 1a). A parallel process would analyze the presence or absence of a target in channels associated with both visual fields simultaneously (see Fig. 1b). A coactive model would combine the information from two parallel channels and make a presence versus absence decision based on that information (Fig. 1c).

Let $f_{L(l)}(t)$ be the density function of the left visual system's detection times when the brightness is low and $f_{L(h)}(t)$ be the left visual system's detection time density when the brightness is high. Likewise, let $f_{R(l)}(t)$ be the right visual system's detection time density when the brightness is low and $f_{R(h)}(t)$ be the right visual system's detection time density when the brightness is high. Let $S_{L(l)}(t)$, $S_{L(h)}(t)$, $S_{R(l)}(t)$, and $S_{R(h)}(t)$ be the survivor functions corresponding to $f_{L(l)}(t)$, $f_{L(h)}(t)$, $f_{R(l)}(t)$, and $f_{R(h)}(t)$, respectively. A survivor function of a random time T is defined as

$$S(t) = P\{\mathrm{T} \geqslant t\} = \int_t^{\infty} f(t')\, dt' = 1 - F(t),$$

$F(t)$ being the cumulative distribution function. The survivor functions are employed rather than the distribution functions because when the minimum time statistic is under study, as in the present investigation, the mathematical expressions are considerably simpler when put in survivor function terms.

Factorial Predictions for Parallel, Serial, and Coactive Models

There are some important assumptions that underpin the application of systems factorial methodology. The most critical of these is that of "selective influence" (e.g., Sternberg, 1969; Townsend & Ashby, 1983). The reader uninterested in the details may proceed to Proposition 1.

There are two aspects of selective influence. The first is simply that an experimental factor affects only a single subprocess. An investigation of what occurs in serial and parallel models when selective influence is violated appears in Townsend & Thomas (1994). We shall assume that selective influence is in force throughout the present study.

The second aspect relates to the level at which a specific factor operates. Since all psychological subprocesses appear to possess variance (i.e., are not deterministic), an experimental variable acting selectively could in principle, affect any one or more of several aspects of a RT distribution. It turns out that the influence must be at a sufficiently strong level to force the underlying architecture to reveal itself through interesting factorial interactions (or lack thereof). A stronger level implies the weaker level but not vice versa. For instance, a factor might simply influence an ordering of the means, as when high intensity speeds up the mean RT. However, mean order is rather weak mathematically. Townsend & Ashby (1978, 1983) and Townsend (1990) have established a hierarchy of ordering relations for statistical distributions and discussed applications in modeling and statistical inference.

A moderate level of ordering is found where one distribution function is always greater than another, for all t. This level implies an ordering of the means but not vice versa (e.g., see Townsend, 1990). Suppose $F_A(t) > F_B(t)$ for all t. Then the survivor functions have the reverse order, $S_B(t) > S_A(t)$ for all t. We shall assume that selective influence holds at this or a slightly stronger level. For instance, on the left, $S_{L(l)}(t) > S_{L(h)}(t)$, with the interpretation that it is always more likely that the information from the high intensity condition has been processed than that the information from the low intensity condition has been processed, for any t. Figure 3a shows a hypothetical example of selective influence operating at the level of survivor function ordering. This implies that the mean processing time for the l (low) condition is longer than that for the h (high) condition, since an ordering in survivor functions implies an ordering in the associated means (Townsend, 1990). Note that S_L is not a function of the R factor or vice versa, in consonance with the selective influence assumption.

A stronger level that will occasionally be useful in the proofs is the assumption that the two process time density functions at the two factor levels cross exactly once and in a non-trivial way. For instance, on the left side, there is exactly one time t_* such that $f_{L(l)}(t_*) = f_{L(h)}(t_*) > 0$, that is,

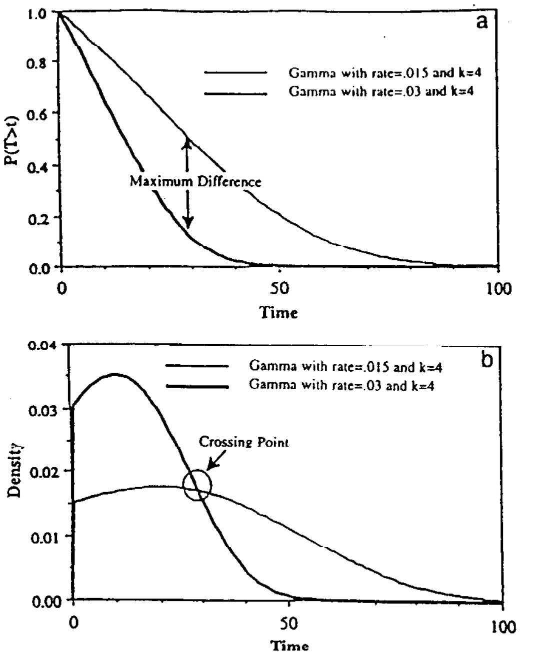

FIG. 3. (a) Two gamma distributed variables exhibit the property of survivor function ordering when the rate is varied, possibly corresponding to an experimental factor influence. (b) The same gamma distributed variables exhibit the even stronger ordering property of a single density crossover, again potentially caused by an experimental factor. See text for more detail.

where the probability density functions for the two stimulus intensities are equal. For $t < t_*$, $f_{L(h)}(t) > f_{L(l)}(t)$ and for $t > t_*$, $f_{L(h)}(t) < f_{L(l)}(t)$. This implies that the survivor functions are ordered, $S_{L(l)}(t) > S_{L(h)}(t)$ for all $t > 0$ (Townsend, 1990; Townsend & Ashby, 1978). Selective influence at the one-point density cross-over level is depicted in Fig. 3b.

Although an ordering in the distribution functions (equivalently, survivor functions) does not imply a single crossover point, it does imply an odd number of non-trivial crossovers (Townsend, 1990). There is accumulating evidence that distribution (and therefore survivor) ordering holds in experiments analogous to the present one (e.g., Ashby, 1982; McGill & Gibbon, 1965; Townsend, 1990; Townsend & Ashby, 1983). Furthermore, it may be that the density single point crossover property may well hold within the range of stable (estimable) data points (Ashby, 1982; Townsend, 1990).

A new powerful tool in systems factorial technology emerges in the form of interactions at the distributional level. As we shall see, survivor functions themselves exhibit distinctive behavior in parallel and serial systems, and also with different stopping rules, under factorial influence. In some cases these qualitatively mirror the mean RT interactions but there are some major surprises where, in fact, the survivor interaction contrasts permit additional tests not seen at the mean interaction contrast level. Even where the qualitative predictions are similar, the distributional predictions are logically more powerful since they generally hold implications for relations among the means but not conversely (Townsend, 1990).

Basically, to produce the survivor interaction contrast, one simply takes the same kind of double difference performed for the mean interaction contrast level, but does it at every time bin where the survivor function is estimated. Theoretically, the models predict a continuous interaction contrast function for every positive time t. Such a curve is approximated by the procedure of subtracting the estimated survivor functions from one another. This function has the constitution

$$S_{ll}(t) - S_{lh}(t) - S_{hl}(t) + S_{hh}(t),$$

where t is the observed RT and the subscripts refer to "left position, low intensity," etc. Basically, the concepts of underadditivity, overadditivity, and additivity apply here just as in the case of the means, except that now any one of these could appear at any particular value of time, t (experimentally, within any observational time bin). For instance, the survivor function interaction contrast will be overadditive at a particular time point, if and only if the above expression is greater than 0 at that time point. These notions are developed rigorously in Appendix A.

In order to avoid technical detail in the main discourse, we adopt the convention in this paper of using the term "proposition" to name the less formal statements of theoretical results in the text that are proven rigorously in the Appendices. The corresponding theorems in the Appendices will always be so designated.

Proposition 1 corresponds to Theorem 1 in Appendix A.

PROPOSITION 1 (Minimum Time Parallel RT Interaction Contrast). *Under general conditions, parallel minimum time process (horse race) models predict overadditivity in both the survivor function and mean* RT *interaction contrasts.*

It should aid the reader to understand the theory and implied methodology if we develop a very special case. Assuming independence, it is easy to prove the result, especially if we momentarily ignore the base time contribution. In fact, the mode of proof adapted by Townsend & Ashby (1983, p. 374) for exhaustive *means* is actually based on an *under*additive interaction of *distribution functions*. The latter is easily converted to the survivor function interaction in self-terminating processing as we now show.

Again, assume self-terminating parallel processing, which in the presence of redundant targets implies a minimum time stopping rule (i.e., a horse race). For the present derivation, recall that the survivor function at time t, for both channels processing simultaneously must require that both channel times be greater than t. Furthermore, when the channels are stochastically independent, that probability is given by the product of the two separate survivor functions. Let interaction contrast of the survivor function in this context be designated IC(SF, ST, t), where SF is the survivor function, ST means self-terminating, and t is the time under consideration. Let $(T_{L(i)}, T_{R(j)}) > t$ represent the event here that both the right and left random times (at respective levels i, j) are greater than t. With these assumptions plus selective influence at the survivor function level in hand we can write the survivor interaction contrast as

$$\begin{aligned}
&\mathrm{IC}(\mathrm{SF}, \mathrm{ST}, t)\\
&\quad = P[(\mathrm{T}_{\mathrm{L(l)}}, \mathrm{T}_{\mathrm{R(l)}}) > t] - P[(\mathrm{T}_{\mathrm{L(l)}}, \mathrm{T}_{\mathrm{R(h)}}) > t]\\
&\qquad - \{P[(\mathrm{T}_{\mathrm{L(h)}}, \mathrm{T}_{\mathrm{R(l)}}) > t] - P[(\mathrm{T}_{\mathrm{L(h)}}, \mathrm{T}_{\mathrm{R(h)}}) > t]\}\\
&\quad = P(\mathrm{T}_{\mathrm{L(l)}} > t)\, P(\mathrm{T}_{\mathrm{R(l)}} > t) - P(\mathrm{T}_{\mathrm{L(l)}} > t)\, P(\mathrm{T}_{\mathrm{R(h)}} > t)\\
&\qquad - [P(\mathrm{T}_{\mathrm{L(h)}} > t)\, P(\mathrm{T}_{\mathrm{R(l)}} > t) - P(\mathrm{T}_{\mathrm{L(h)}} > t)\, P(\mathrm{T}_{\mathrm{R(h)}} > t)]\\
&\quad = S_{\mathrm{L(l)}}(t)\, S_{\mathrm{R(l)}}(t) - S_{\mathrm{L(l)}}(t)\, S_{\mathrm{R(h)}}(t)\\
&\qquad - [S_{\mathrm{L(h)}}(t)\, S_{\mathrm{R(l)}}(t) - S_{\mathrm{L(h)}}(t)\, S_{\mathrm{R(h)}}(t)]\\
&\quad = [S_{\mathrm{L(l)}}(t) - S_{\mathrm{L(h)}}(t)][S_{\mathrm{R(l)}}(t) - S_{\mathrm{R(h)}}(t)].
\end{aligned}$$

Because of the action of selective influence on the survivor functions, $S_{\mathrm{L(l)}}(t) > S_{\mathrm{L(h)}}(t)$ and $S_{\mathrm{R(l)}}(t) > S_{\mathrm{R(h)}}(t)$ for all $t > 0$, proving that IC(SF, ST, t) > 0. Furthermore, if one could neglect the base time, then one simply integrates IC(SF, ST, t) over all positive t to obtain the analogous result for mean processing time (special case of Theorem 1B).

Next, it has been known for some time that parallel exhaustive processing typically produces mean interaction contrast underadditivity (Schweickert, 1978; Townsend & Ashby, 1983; Townsend, 1984). The following proposition shows that this underadditivity is also characteristic of the survivor interaction contrast.

Proposition 2 (Parallel Exhaustive RT Interaction Contrast). *Under general conditions, parallel exhaustive process models predict underadditivity in both the survivor and mean* RT *interaction contrasts.*

Proposition 2 ensues immediately from combining of Theorems 1 and 2 in Appendix A.

It was learned above that in the context of parallel processing, interactions at the distributional level behave in the same qualitative fashion as do the means. Now consider serial processing. When processing is self-terminating, serial models follow suit, with both means and survivor functions exhibiting additivity (zero interaction). Interestingly, this is *not* the case for serial *exhaustive* processing. In fact, under certain conditions that ensure selective influence, we later show that for short RTs, the factorial interaction is negative for "exhaustive" survivor functions. Of course, it is well known that additivity holds for the exhaustive serial mean RTs.

Proposition 3 corresponds directly to Theorem 3 in Appendix A.

Proposition 3 (Serial Minimum Time RT Interaction Contrast). *Under general conditions, the minimum-time serial* RT *survivor interaction contrast and mean interaction contrast are predicted to be zero.*

The serial exhaustive models are interesting in that they must predict regions of positive and negative interaction contrast. It will be observed that, although both serial self-terminating and serial exhaustive models predict mean interaction contrast to be zero, the survivor interaction contrast functions are distinct. While serial minimum time models predict a survivor interaction contrast function that is always zero (Proposition 3), the serial exhaustive models predict a survivor function interaction contrast that starts out negative and becomes positive later on.

Proposition 4 (Serial Exhaustive RT Interaction Contrast). *Under general conditions, the serial exhaustive models predict survivor interaction contrast functions that are negative for small times t and later become positive. The serial exhaustive mean* RT *interaction contrast is predicted to be zero.*

Theorem 4 in Appendix A corresponds to Proposition 4. To get an idea of what happens in the intriguing case of survivor function underadditivity with exhaustive serial processing, let us consider a special case of the fully independent situation of the corollary to Theorem 4 in Appendix A. Let t_{B} stand for the base (i.e., residual) times associated with extraneous processes and, of course, L refers to left, R to right, and l is low and h is high. Basically, the survivor interaction contrast for exhaustive serial processing involves taking the double difference (across the two factors) of the overall survivor function for each condition (i.e., l on left, h on right, etc.). In turn, the overall survivor function is composed of a convolution of the base time distribution, the left distribution, and the right distribution. The result, in the independent case, is the following survivor function interaction contrast:

$$\begin{aligned}
&\mathrm{IC}(\mathrm{SF}, \mathrm{EXH}, t)\\
&\quad = \int_{t_{\mathrm{L}}=0}^{t} \int_{t_{\mathrm{B}}=0}^{t-t_{\mathrm{L}}} [S_{\mathrm{R(l)}}(t - t_{\mathrm{L}} - t_{\mathrm{B}}) - S_{\mathrm{R(h)}}(t - t_{\mathrm{L}} - t_{\mathrm{B}})]\\
&\qquad \times [f_{\mathrm{L(l)}}(t_{\mathrm{L}}) - f_{\mathrm{L(h)}}(t_{\mathrm{L}})]\, f_{\mathrm{B}}(t_{\mathrm{B}})\, dt_{\mathrm{B}}\, dt_{\mathrm{L}}.
\end{aligned}$$

Now observe that the first term under the integral sign is always positive, due to direct selective influence at the level of survivor ordering. Further, if selective influence operates at the level of the density one point crossing property, then the second term will be negative for $t < t_*$, where t_* equals the density crossover point. However, even if the densities cross more than once, it must be the case that $f_{L(h)}(t) > f_{L(l)}(t)$ for small t (recall that for survivor function ordering to occur, the densities must cross not at all or must cross an odd number of times; Townsend, 1990). Otherwise the distribution functions would not be ordered as assumed, so the result still holds.

The class of coactive models has been a little fuzzy around the edges. The most primitive and inchoate idea seems to be that somehow the two original channels (usually taken to be arranged in parallel) give rise to something new over and above the separate information that permits faster responses than what is expected from separate activations. The "separateness" of activation of parallel channels was given operational definition by Miller (e.g., 1982) in terms of an inequality to be discussed further later. Realizations of the coactive concept that are more theoretically oriented have appeared over the last few years.

One type of model that seems natural assumes that each parallel channel can be represented as a separate counting mechanism accruing information from its input source. If a certain criterion number of counts is accrued by time t, then a "detection" occurs and a response can be made. When either of the two channels can make a detection, we have a parallel model. If the counts are sent to a final common channel, we have instead created a special kind of coactive model. We call such models "channel summation models" for obvious reasons.

The exact specification of particular varieties of channel summation model will differ somewhat depending on the questions asked and the generality at which we are able to prove our theorems. When the counting processes are governed by independent Poisson distributions, we produce the special case known as the "neural summation model" studied by Schwarz (1989) and Colonius and colleagues (Colonius, 1990a; Diederich & Colonius, 1991). In keeping with the present notation, we shall refer to such models as "Poisson channel summation models."

Along with data that are associated with capacity to be developed in the next section, interaction contrast predictions aid in discriminating parallel race models from channel summation models. In fact, we shall see that at the mean RT interaction contrast level, channel summation models make the same qualitative prediction as parallel race models. However, at the survivor interaction contrast level, they make distinctive predictions (see also Theorem 5 in Appendix A).

PROPOSITION 5 (Poisson Channel Summation Model RT Interaction Contrast). *Under general conditions, the Poisson channel summation (coactive) model predicts survivor interaction contrast functions that start out negative and later turn positive. The mean interaction contrasts are always positive.*

TABLE 2

Summary of Mean and Survivor Function Interaction Contrast Propositions

	Mean interaction contrast	Survivor interaction contrast
S-T[a] (min. time) parallel	Positive	Positive
EXH[b] (max. time) parallel	Negative	Negative
S-T (min. time) serial	Zero	Zero
EXH (max. time) serial	Zero	Neg → pos
Channel summation	Positive	Neg → pos

[a] S-T, self-terminating.
[b] EXH, exhaustive.

Thus, the surprising result is obtained that an early negative survivor interaction contrast is predicted by this model just as it is by the exhaustive serial class as noted above. However, it can be shown that the amount of negative area under the survivor interaction contrast functions must equal the amount of positive area for serial exhaustive models whereas the negativity in channel summation models is usually much smaller than the amount of positivity. At present we lack a solid intuition for the positive vs negative interaction contrast predictions by the various models. It is one of the benefits of mathematical modeling that decisive tests often arise from the mechanisms of theorem production rather than out of (sometimes faulty) intuition.

Table 2 summarizes the interaction contrast predictions for the various models.

Discussion of Technical Conditions for Interaction Contrast Propositions

The predictions associated with the above propositions are parameter and distribution free, meaning that they do not depend on any specific stochastic process being assumed to underlie processing. They also allow for the inclusion of a residual RT variable, that is, the "extra" time contributed by the sensory, motor, and mental processes not under study. The residual time is often called the "base time."

Architectures other than serial or parallel can, in some cases, mimic the factorial predictions of these. For instance, a bare Wheatstone bridge predicts underadditivity in an exhaustive processing situation (Schweickert, 1978; Schweickert & Townsend, 1989). However, these more complex networks do not seem viable in the present very simple perceptual situation, where one dot may or may not be presented dichoptically to each eye. In fact, the physiology of the visual system, together with stimuli impinging on non-corresponding parts of the two eyes, suggests at least an early parallel phase of processing. However, both the architecture and the stopping rule (self-terminating vs exhaustive) are empirical questions.

Next, whether processing in separate subsystems or on separate items (whether serial, parallel, etc.) is statistically independent (often called stochastic independence) is of interest in its own right. The answer to the question of statistical independence also impacts heavily on the likelihood that the assumption of selective influence is justified. Selective influence implies that each experimental factor influences exactly one psychological process or subsystem. Selective influence is in some ways the most critical aspect of the technology, for without it, we presently have no control or prediction. As shown previously (Townsend, 1984; Townsend & Ashby, 1983), selective influence can be disrupted either by direct means or indirectly through statistical dependencies. Statistical independence is logically independent of selective influence in the original additive factor method (Sternberg, 1969; Townsend & Ashby, 1983) and in the structures covered by the trichotomy method (Townsend & Schweickert, 1989). However, recent mathematical work shows that statistically dependent parallel and serial processes can predict any of the three major types of interaction (overadditivity, underadditivity, and ordinary additivity), through indirect nonselective influence (Townsend & Thomas, 1994). The latter paper investigates such matters in detail and provides clues for detecting a failure of selective influence. Dzhafarov (1992) proposes a class of models where there is total dependence among sequentially arranged processes and Dzhafarov & Schweickert (1994) investigate additive factorial results at the distributional level in such systems.

Our present theorems on interactions predicted by parallel and serial systems under self-terminating or exhaustive processing permit statistical dependencies between the parallel channels themselves, as long as certain conditions are met. The present conditions, on dependencies, although not identical to those of Townsend & Schweickert (1989), serve an analogous purpose.

Our most general factorial prediction theorems also allow stochastic dependence between the processing times and the base times that add to the processing times to make up the observable RTs. Most modeling effects have postulated stochastic independence of the base time and processing time (see Luce, 1986; Townsend & Ashby, 1983). The paper by Dzhafarov (1992) mentioned just above investigates models that assume a total dependence of base time on processing time in a serial architecture. Much research is needed on this important aspect of processing.

Parallel Models' and Coactive Models' Capacity Predictions

It is well known that a redundant target condition can lead to shorter RTs even if the processing is independent parallel and both channels process at the rate each did when acting alone. A full understanding of the implications of this kind of result, as well as the theoretical tests employed below, calls for a brief consideration of the notion of capacity in RT process models. The discussion follows the approach to capacity developed by Townsend and colleagues (e.g., Townsend, 1972; 1974; Snodgrass & Townsend, 1980; Townsend & Ashby, 1978; 1983) and readers already versed in this literature may wish to move directly to the subsequent propositions. Serial models are not pursued further because the interaction contrast analyses developed above decisively falsify them in the present experiments.

Previous work has demonstrated that the effects of load, for instance, number of items input to the overall system, can depend on architecture (e.g., serial vs parallel), stochastic dependence, and stopping rule (e.g., Colonius & Vorberg, 1994) and, obviously, individual item or channel processing capacity (e.g., Townsend, 1972, 1974). We expand that knowledge here, both to redundant-target parallel processing and to the new coactive theories.

This approach to capacity is rigorous and process oriented in terms of its detailed mathematical treatment and predictions. Nevertheless, it does not imply a complete theory of capacity in that it is mute with regard to the ultimate reservoir of process capacity and relationships among simultaneous tasks (for example), as found in the proposed model, say, of Kantowitz & Knight (1974). It does seem to fall within the general scope of the scenarios of Kahneman (1973) or Norman & Bobrow (1975) and Sanders (1983) and is susceptible to further extension.

There are two closely related and mutually compatible ways of employing capacity (Townsend, 1974; Townsend & Ashby, 1983, Chapter 4). The first applies to some designated unit of analyses, such as an item or features of an item. Thus, in a parallel model, the capacity of an item processor may decrease as the total number of items undergoing processing increases. Another meaning of the word relates to some observable measure in a particular paradigm or condition. An example is mean RT when a serial processor can self-terminate on a target present trial. It should be evident that this level is coarser than that of an individual item and in fact, assumptions about capacity at, say, the item level force predictions about the more macroscopic level of observed mean RT.

The next thing to notice is that RT capacity effects, for instance, increase of RT as load is incremented, may be due to real capacity effects at the item level or simply statistical effects. For instance, an independent, unlimited capacity (no capacity decrements at the individual item level when load increases) parallel model predicts increases in mean RT when it must process all items exhaustively. This increase in RT is caused by purely statistical effects (Townsend, 1974; Townsend & Ashby, 1983, pp. 83, 84).

Now consider what happens in an experiment like the present one, when we move from a single dot type of trial to a double dot type of trial. In the present circumstances, the completion of either item immediately conveys "at least one target present" information and thereby permits a correct "yes" response. The result is that when the system is parallel and can terminate as soon as the first signal (dot) is processed, a statistical "race" advantage (rather than a disadvantage as with exhaustive processing, for example on "no target" trials) is present. The difference of the exhaustive (wait until the last horse finishes the race) from minimum time completion (stop as soon as the first horse finishes) is that in the former, the statistical effects operate in the same direction as any limited capacity effects produced from an increase in processing load. In the minimum time race situation, the increase in number of items or operating channels pits a statistical advantage (shorter response times) against a possible limitation in capacity effect (longer response times). Hence, in any redundant targets paradigm where positive trials (i.e., trials with targets present) present only targets with no distractors, the question of capacity inevitably enters in comparisons of fewer targets and more targets, RTs.

To summarize, the experimenter must be on the lookout for a possible tradeoff between the benefits due to redundancy when the system can self-terminate and the potential detriments due to capacity limitations at the individual item level. When the parallel architecture is unlimited capacity with statistical independence of its channel distributions, processing efficiency on each channel is unimpaired by the increase in load (one to two targets). When a minimum time stopping rule (the race) is imposed on such a model the predicted improvement in RT is known, somewhat inappropriately, as "probability summation" as noted earlier (e.g., Meijers & Eijkman, 1977; Raab, 1961) and is totally a statistical effect. As noted, this is the opposite of what happens when one waits for the maximum of two independent parallel unlimited-capacity channels to finish, rather than either alone (e.g., Egeth, 1966; Townsend, 1974). As will be seen shortly, coactivation may also be convincingly portrayed in terms of capacity mechanisms.

A natural way to measure capacity in RT environments is through the hazard function or a close relative which we develop below. The hazard function completely determines a probability distribution and vice versa (e.g., Townsend & Ashby, 1983, pp. 26, 27). The hazard function $h(t)$ may sometimes be interpreted in terms of the concept of power and $H(t)$, the integral of the hazard function, as energy (i.e., summed power, or work done across time; Townsend & Ashby, 1983, pp. 76–80). The hazard function can also be thought of as a conditional density function. That is, given that an item is not finished by time t, $h(t)=f(t)/S(t)$ is proportional to the probability that it finishes in the next instant. Now $h(t)$, considered to be the hazard function for a single item in any model, may change as a function of the number of items to be processed. Again, this possibility of a decrease in $h(t)$ or $H(t)$ is especially viable in parallel channels. Hence, it can be seen that the hazard function, or its integral, forms a genuine measure of capacity within process models. As an example, the exponential rates in exponential or gamma distributed parallel models are special case instances of hazard functions that reflect capacity relative to load pressures (e.g., Townsend & Ashby, 1983). Let $h(t)$ be the hazard function. Then we can express the survivor function as $S(t)=\exp[-\int_0^t h(t')\,dt']$, (see Townsend & Ashby, 1983, Eq. 3.5). Similarly, $H(t)=\int_0^t h(t')\,dt'$, the integrated hazard function.

In order to develop a measure of capacity, a good benchmark is the unlimited capacity independent parallel model. Under stochastic independent parallel processing, unlimited capacity is in force if exactly the same parallel left and right processing time distributions apply when both items (e.g., dots) are present, as when only one is present. Then the minimum processing time survivor function is

$$S_{L,R}(t)=S_L(t)\cdot S_R(t).$$

The latter equation then yields an expression in terms of the integrated hazard functions,

$$\begin{aligned}S_{L,R}&=\exp(-H_{L,R}(t))=S_L(t)\cdot S_R(t)\\&=\exp[-H_L(t)]\cdot\exp[-H_R(t)],\end{aligned}$$

and it can be seen by taking the natural logarithm on both sides of the above expression that

$$H_{L,R}(t)=H_L(t)+H_R(t).$$

The fact that $H_{L,R}$ equals the sum of the two separate capacity terms, each of which applies when only its item is present, is equivalent to *unlimited capacity* independent parallel processing. That is, the processing capacity in each channel is exactly as good as but not better than it was when only a single dot was present. However, capacity could be reduced with the presence of two dots rather than one. In a *limited capacity* model,

$$H_{L,R}(t)<H_L(t)+H_R(t).$$

Thus, the combined processing capacity when both dots are present is less than the sum of the two individual capacities, the latter being represented by the respective integrated hazard functions.

Now, in pursuing the parallel unlimited capacity model as a benchmark, we must allow for the possibility that the double dot condition could produce better (i.e., faster) performance than that expected from a parallel unlimited capacity race. Hence, in any model that is able to produce *super capacity* when compared with parallel unlimited capacity processing we would have the following outcome:

$$H_{\mathrm{L,R}}(t) > H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t).$$

This leads us to define the *capacity coefficient* as

$$C(t) = \frac{H_{\mathrm{L,R}}(t)}{H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t)}, \quad t > 0.$$

That is, the capacity coefficient measures capacity (integrated hazard function) in the combined dot condition against (i.e., takes the ratio) that of the sum of the two single dot conditions. The coefficient of capacity is greater than, equal to, or less than 1 according to whether capacity is super, unlimited, or limited, respectively, at that particular value of t. The fact that $C(t)$ measures capacity at all values of time means that it is a relatively fine-grained index with regard to processing efficiency. We may then write

$$H_{\mathrm{L,R}}(t) = C(t)[H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t)]$$

for various theoretical purposes.

Reiterating, the integrated hazard function of the *unlimited* capacity parallel independent model is equal to the sum of the individual hazard functions from the single signal conditions and $C(t) = 1$ for all times t. In the *limited* capacity case, the integrated hazard function of the minimum completion time is smaller than the sum of the individual hazard functions and $C(t) < 1$. In the *super* capacity case, the hazard function of the minimum completion time is greater than the sum of the individual hazard functions and $C(t) > 1$. In special cases $C(t)$ can be a constant, C, as found in exponential parallel models (Townsend & Ashby, 1983, Chapter 4).

The above relationships permit us to express the key survivor functions in a manner that will facilitate testing critical predictions. Thus, where $C(t)$ is the capacity coefficient,

$$\begin{aligned} S_{\mathrm{L,R}}(t) &= e^{-H_{\mathrm{L,R}}(t)} = e^{-C(t)(H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t))} \\ &= (e^{-H_{\mathrm{L}}(t)} \cdot e^{-H_{\mathrm{R}}(t)})^{C(t)} = (S_{\mathrm{L}}(t) \cdot S_{\mathrm{R}}(t))^{C(t)}. \end{aligned}$$

In the unlimited-capacity case, $S_{\mathrm{L,R}}(t) = S_{\mathrm{L}}(t) \cdot S_{\mathrm{R}}(t)$, because $C(t) = 1$; in the limited-capacity case, $S_{\mathrm{L,R}}(t) > S_{\mathrm{L}}(t) \cdot S_{\mathrm{R}}(t)$, because $C(t) < 1$; and in the supercapacity case, $S_{\mathrm{L,R}}(t) < S_{\mathrm{L}}(t) \cdot S_{\mathrm{R}}(t)$, because $C(t) > 1$.

As noted earlier, Miller (1982) proposed a now famous inequality based on a fundamental property of probability distributions to test a class of race models.[1] We now present this test within the above theoretical purview. The test is constructed as follows on the distribution functions F. Recall that $F(t) = 1 - S(t)$ and let $F_{\mathrm{L,R}}(t) = P(\mathrm{T}_{\mathrm{L,R}} \leqslant t)$ be the redundant target (both signals present) distribution function. Integrating over the right side density, call $P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{L}} \leqslant t)$ the marginal left side probability distribution in the redundant condition and similarly for $P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{R}} \leqslant t)$ on the right. Next, let $P_{\mathrm{L}}(\mathrm{T}_{\mathrm{L}} \leqslant t) = F_{\mathrm{L}}(t)$ be the single stimulus (left side alone) probability distribution and likewise for $P_{\mathrm{R}}(\mathrm{T}_{\mathrm{R}} \leqslant t) = F_{\mathrm{R}}(t)$. Then with a minimum-time stopping rule,

$$\begin{aligned} F_{\mathrm{L,R}}(t) &= P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{L}} \leqslant t \text{ or } \mathrm{T}_{\mathrm{R}} \leqslant t) \\ &= P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{L}} \leqslant t) + P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{R}} \leqslant t) \\ &\quad - P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{L}} \leqslant t, \mathrm{T}_{\mathrm{R}} \leqslant t) \\ &\leqslant P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{L}} \leqslant t) + P_{\mathrm{L,R}}(\mathrm{T}_{\mathrm{R}} \leqslant t) \\ &\leqslant P_{\mathrm{L}}(\mathrm{T}_{\mathrm{L}} \leqslant t) + P_{\mathrm{R}}(\mathrm{T}_{\mathrm{R}} \leqslant t) \\ &= F_{\mathrm{L}}(t) + F_{\mathrm{R}}(t), \end{aligned}$$

so

$$F_{\mathrm{L,R}}(t) \leqslant F_{\mathrm{L}}(t) + F_{\mathrm{R}}(t).$$

Note that though independence is not necessitated for this inequality, there *are* implicit assumptions with regard to capacity: Namely, the inequality is based upon the assumption that the marginal distributions from the redundant target condition are less than or equal to the respective distribution functions from the left-only and the right-only conditions, as is pointed out by Ulrich & Giray (1986) and by Ashby & Townsend (1986; see also Colonius, 1990b). This is fundamentally an assumption that capacity in any race model would be at best unlimited capacity but could be limited capacity. This assumption seems relatively innocuous because we do not in general expect parallel race models to be super capacity. However, we shall see later that parallel models that permit interchannel cross talk could produce super capacity. Miller (1982) found many violations of this inequality and rejected all "race" models. Although he did not draw conclusions with regard to capacity, violation of his inequality implies super capacity in the present theory.

There is another bound of interest, although it is less well known in the literature. We call it the Grice bound because Grice and his colleagues appear to have been the first to employ it (see Grice, Canham, & Gwynne, 1984). Basically it establishes a bound on limited capacity, as opposed to Miller's bound on super capacity. The bound is given by $\mathrm{MAX}\{F_{\mathrm{L}}(t), F_{\mathrm{R}}(t)\}$; note that when capacity is parallel, independent, and unlimited, this will always be less than

[1] The inequality that follows was, of course, already well known in probability theory ("Boole's inequality"). However, many fields permit the attachment to the name of a technique the name of the scientist who first discovers a valuable use for it.

$F_{L,R}(t) = F_L(t) + F_R(t) - F_L(t)\, F_R(t)$. Putting the two inequalities into a single expression we arrive at

$$\mathrm{MAX}\{F_L(t), F_R(t)\} \leqslant F_{L,R}(t) \leqslant F_L(t) + F_R(t).$$

The working hypothesis in the literature has been that if processing (read "activation" in many accounts) is separate rather than "coactive" and implicitly not too super or limited in capacity, then the observed redundant target distribution function should lie within the stated bounds. This would be the case in independent unlimited capacity parallel processing, for example. We now present the survivor function version of the Miller and Grice bounds.

It is critical to note at this juncture that the two bounds reverse themselves in moving from the distribution form to the survivor form. The larger the survivor function, the slower the processing time. Subsequently, we will always mean the Grice survivor function bound when we refer to the "upper bound" and to the Miller survivor function bound when we refer to the "lower bound." The upper, Grice bound then becomes the "lid" on moderately limited capacity: any number greater is *very* limited in capacity. The Miller bound becomes the "lid" on moderately super capacity and any number less indicates *extreme* super capacity. This latter statement will become clearer as we proceed by substituting 1-$S(t)$ for $F(t)$ in every instance in the inequalities above. The survivor bounds are then

$$\mathrm{MIN}\{S_L(t), S_R(t)\} \geqslant S_{L,R}(t) \geqslant S_L(t) + S_R(t) - 1,$$

where $S_{L,R}(t)$ is the survivor function when both the right and the left signals are presented. We now employ the survivor functions, $S_L(t)$ and $S_R(t)$, from the single target conditions as a "base line" with which to measure capacity in moving from the single to redundant target condition.

This strategy permits the bounds to be given a natural interpretation in terms of the capacity coefficient $C(t)$. First, consider the Grice upper bound of the inequality and how it can be violated. Suppose $S_L(t) = \mathrm{MIN}\{S_L(t), S_R(t)\}$, from the single target conditions. Then the following inequality results when the observed survivor function of the redundant signal condition $S_{L,R}(t) = [S_L(t) \times S_R(t)]^{C(t)}$ is greater (implying violation) than

$$S_L(t) = \mathrm{MIN}\{S_L(t), S_R(t)\}, \qquad \text{that is,}$$

$$S_L(t) < [S_L(t) \times S_R(t)]^{C(t)}.$$

Taking the logarithm and solving for $C(t)$, we obtain

$$C(t) < [\ln S_L(t)]/[\ln S_L(t) + \ln S_R(t)].$$

In terms of integrated hazard functions,

$$C(t) < \frac{H_L(t)}{H_L(t) + H_R(t)}.$$

But note that $\ln\{\mathrm{MIN}[S_L(t), S_R(t)]\} = -\mathrm{MAX}[H_R(t), H_L(t)]$. Hence, to violate the Grice inequality, $C(t)$ has to be less than the largest of the two integrated hazard functions divided by the sum. Thus, violation of the Grice bound implies limited capacity of a rather strong degree. In general, the capacity bound is close to the level of *fixed capacity*, defined by $H_{L,R}(t) = \frac{1}{2}[H_L(t) + H_R(t)]$ (Townsend & Ashby, 1978), because

$$\tfrac{1}{2}[H_L(t) + H_R(t)] \leqslant \mathrm{MAX}[H_L(t), H_R(t)],$$

with equality (and fixed capacity) holding only in the event that $H_L(t) = H_R(t)$. Observe that the Grice bound is on the right-hand side. When the left- and right-hand integrated hazard functions are identical, the Grice bound becomes equivalent to fixed capacity, but otherwise is a bit greater. Hence if $C(t) < \frac{1}{2}$, the system is deemed to be of *very* limited capacity at those values of t where this occurs.

Next, consider the Miller lower bound of the inequality, in terms of survivor functions. Here, we are interested in how much extra capacity is required to violate the lower bound of the inequality. First, we have

$$S_{L,R}(t) = \{S_L(t) \times S_R(t)\}^{C(t)} \geqslant S_L(t) + S_R(t) - 1.$$

A violation of the inequality then implies the reverse:

$$\{S_L(t) \times S_R(t)\}^{C(t)} < S_L(t) + S_R(t) - 1.$$

This inequality leads to Proposition 6, which we state in a more intuitive form here and technically in Theorem 6 (Appendix B).

Proposition 6 (The Relation of Miller's Inequality to Capacity). (i) *If Miller's inequality is violated at some value of time t, then the system is super capacity at that time, t.*

(ii) *If the system is everywhere super capacity, then Miller's inequality will be violated for some values of t.*

Proposition 6 means that using parallel processing as a benchmark, a violation of Miller's inequality implies an incidence of super capacity whatever the underlying system that produces the violation. The super capacity, for instance, is produced automatically by a large and natural variety of coactivation models, as we shall see in the following two propositions. It is also important to recognize that although super capacity and Miller's capacity are related through Proposition 6, they are not identical concepts. If they were, the mathematics would demonstrate that fact.

We next introduce an important variety of parallel "race" model (i.e., with minimum time stopping rule). Suppose each separate channel acquires information supporting the presence of a target by way of a series of randomly distributed counts (see, e.g., Townsend & Ashby, 1983, Chapter 9). If a criterion number of counts is recorded in either channel a "detection" occurs. In the present context, these notions would normally give rise to a special class of independent parallel race model falling within the general class of parallel models treated throughout this paper.

A highly useful property of counting models is that the probability function on the waiting time until some criterion number of counts occurs is intuitively related to the probability function on the counting process. Let (N_i, t) be the random number of counts accrued on channel i by time t. Then $P(\mathrm{T}_i > t \mid \text{Criterion} = k_i) = P[\mathrm{N}_i < k_i, t]$, the probability that processing is not yet completed by time t in channel i equals the probability that the summed counts by time t is less than the criterion required for a "detection" decision.

In fact, with independence assumed, we would have the parallel counting model formula for the redundant targets condition (omitting the base time),

$$S_{\mathrm{L,R}}(t) = P_{\mathrm{L}}(\mathrm{N_L} < k_{\mathrm{L}}, t)\, P_{\mathrm{R}}(\mathrm{N_R} < k_{\mathrm{R}}, t),$$

where $\mathrm{N_L}$ is the random number of counts and similarly for $\mathrm{N_R}$ and k_{L} is the criterion number of counts that triggers a detection decision on the left and analogously for k_{R} on the right. Of course, in a general formulation, we would have to include notation to indicate that the separate counting processes may depend on the load, that is, capacity may be affected where two signals rather than one are presented.

Now, suppose instead of maintaining separate count totals, the information from the two channels is funneled into a common outlet, which serves to sum the outputs of the two parallel processors. We thereby generate the channel summation model met in the earlier sections on factorial interactions.[2] As noted earlier, a natural special case based on Poisson counts was presented by Schwarz (1989) and has already started to receive additional theoretical and experimental attention (Colonius, 1990a; Diederich & Colonius, 1991).[3] Miller (1982) referred to models which could predict a violation of his inequality as "coactivation models."

The channel summation model may be thought of as a specific type of coactivation model (cf. Eriksen, Goettl, St. James, & Fournier, 1989; Meijers & Eijkman, 1977). The idea in the channel summation model is that if the common final channel acquires a criterion number of counts, then it causes a "detect" response. Again, call the independent unlimited capacity processes for the left and the right $\mathrm{N_L}$ and $\mathrm{N_R}$. Then the across-channel summed counting process is expressed

$$\mathrm{N_{sum}} = \mathrm{N_L} + \mathrm{N_R}.$$

If $\mathrm{W}_{\mathrm{sum},k}$ is the random variable representing the waiting time until the sum of the inputs from the two channels hits k, then the relation may be stated in terms of its cumulative distribution function as

$$P\{\mathrm{W}_{\mathrm{sum},k} \leqslant t\} = P\{\mathrm{N_{sum}} \geqslant k, t\} = P\{\mathrm{N_L} + \mathrm{N_R} \geqslant k, t\}.$$

This is the probability that processing is completed by time t for the summation model.

The specific Poisson channel summation model of Schwarz (1989) and others has received initial testing (Colonius, 1990a; Diederich & Colonius, 1991) as noted above. However, it has never been shown to always be super capacity, to always violate Miller's (1982) inequality, or to always satisfy the Grice, Canham, & Boroughs (1984) inequality. We now investigate such questions for a considerably more general class of models which includes the Poisson channel summation model in a special case. It is of significant interest to know whether important qualitative data features characterize entire classes of models as opposed to specific cases of such models.

To summarize the theoretical issues:

(a) Are channel summation models generally truly super capacity, that is, less limited in capacity than parallel independent unlimited capacity models?

(b) Do channel summation models predict violation of the Miller inequality and satisfaction of the Grice, Canham, & Boroughs (1984) bound?

The next two propositions and their related Theorems (Appendix B) address these two questions.

The count criteria k play an important role in the parallel counting horse race models as well as in the counting channel summation models. Without the added information that might arise from speed–accuracy tradeoff or ancillary data, k simply helps, along with the rate parameters, to determine the overall speed. Miller's inequality involves a comparison of single channel and combined RTs. Hence, we found it necessary in the following propositions to assume that the same criterion is employed in all three cases: the left and right parallel "alone" channels as well as the channel summation channel in the double signal condition. In addition, it is supposed that the likelihood that a specific channel has reached any particular criterion by an arbitrary time t, is the same for that channel when both channels are active as

---

[2] This type of processing has often been referred to as "neural summation." We prefer the more neutral term "channel summation."

[3] Schwarz (1989) reports in a footnote information that Miller and Ulrich had also previously worked on such models. Schwarz's appears to be the first published account.

when it alone is operating (i.e., with only a single target presented to that channel). Thus, it is clear that in order to compare the behavior of the independent unlimited capacity parallel model with the channel summation model, these individual channel distributions should be the same. These assumptions will be captured by requiring that the channel summation model's parallel channels will "have the same channel distributions" as an unlimited capacity (ordinary) parallel race model, in the statement of Proposition 7 below.

Proposition 7 indicates that independent channel summation models are indeed super capacity because their measured capacity is greater than that of the unlimited capacity independent parallel model, at the distributional level. Its concomitant is Theorem 7 in Appendix B.

PROPOSITION 7 (The Capacity of Channel Summation Models). *Consider an independent channel summation model based on the same channel distributions as an independent unlimited capacity parallel counting model for all counts k and times t. Then the channel summation model is always of greater capacity than its parallel counterpart. That is, the channel summation model is super capacity for all times t,*

$$\mathrm{MIN}[S_L(t), S_R(t)] \geqslant S_L(t) \times S_R(t) \geqslant S_{CS}(t),$$

where $S_{CS}(t)$ stands for any such channel summation model survivor function.

Note that the independent parallel model, with preserved channel inputs, was employed on purpose in Proposition 7 since it serves as the gauge of super capacity. Hence, if capacity is demonstrated by data to be super capacity, one possible explanatory class of models is that based on independent channels, with channel summation. Recall from Proposition 6 that if a system is super capacity for all time t, then it will violate Miller's bound for some times t. Proposition 7 then implies that the independent channel summation model must produce violations of Miller's inequality. The next proposition shows that even channel summation models with interchannel dependencies, and assuming capacity is not significantly reduced in the load increase, must predict violations of Miller's inequality. This corresponds to Theorem 8 in Appendix B.

PROPOSITION 8 (Channel Summation Models and Miller's Inequality). *Under relatively general conditions, including the stipulation that marginal channel distributions are preserved in going from a single to double target condition, a (possibly dependent) channel summation model predicts violations of Miller's inequality for some times t.*

Obviously, if the marginal channel distributions become larger (an unlikely individual channel super capacity result), violations still occur. Basically, only if the channel capacity is severely degraded in going from one to two targets can the inequality be satisfied.

This generalization of Miller-inequality violations to dependent channel summation models *cannot* be derived from Propositions 6 and 7. This is extremely important because under similar stipulations about what happens to the individual channels when moving from single to double targets, interchannel dependencies can make a regular parallel model appear to be limited capacity. That is, even though the marginal channel probability distributions are preserved in going to the double target condition, observable capacity can still be degraded in such models. However, as we observed earlier, it cannot be so degraded as to violate the Grice inequality, even in ordinary parallel race models. In contrast, channel summation models preserving the individual channel (marginal) distributions (or with them larger) must be sufficiently super capacity as to predict violations of Miller's inequality.

A related point is that it is relatively obvious that a model might be somewhere super capacity (as in Proposition 7), but not enough so to predict a violation of Miller's inequality (Proposition 8). But it might be thought that violation of Miller's inequality (Proposition 8) would force the conclusion of super capacity (Proposition 7), since independent channel summation models (Proposition 7) are a special case of general channel summation models (Proposition 8). This is in fact true for the time durations for which the violation is predicted. However, it cannot be logically inferred from a violation of the inequality that the unlimited capacity parallel model is always (i.e., for all k, t) more limited in capacity than the channel summation model. Thus, Proposition 7 cannot be deduced from Proposition 8 and is of independent interest.

One other implication is that capacity effects of going from one to two targets can force coactive models to act in a limited capacity fashion, for instance, like a limited capacity parallel or even a serial model. This is, of course, reminiscent of what can happen in parallel processing when the load is increased. Coactivation provides for a more dramatic augmentation of capacity via consolidation of activation than do even parallel race alternatives. Hence, capacity degradation must be much more severe in the case of coactivation as compared with, say, parallel race processing, to bring speeds into any given slower range.

To sum up the results of this section: 1. Super capacity is associated with capacity index $C(t) > 1$ at some values, at least, of t. 2. Consistent super capacity, indicated by $C(t) > 1$ for all t, implies a violation of Miller's inequality. 3. Violation of Miller's inequality implies super capacity at the values of t where the violation occurs. 4. Violation of Grice's inequality implicates very limited capacity. 5. Independent channel summation models that preserve the marginal distributions (as do unlimited capacity (independent) parallel models) are always super capacity. 6. General channel summation models that preserve the marginal

TABLE 3

Summary of Capacity and Coactivation Propositions

Supercapacity $\leftrightarrow C(t) > 1$
(Definition in terms of independent parallel model)

Violation of MI[a] at $t = t^* \rightarrow$ supercapacity at $t = t^*$

Supercapacity for all $t > 0 \rightarrow$ violation of MI for some t

ICSM[b] with invariant channels $\rightarrow$ supercapacity for all times t

GCSM[c] with invariant channels $\rightarrow$ violation of MI for some t

[a] MI, Miller's inequality.
[b] ICSM, independent channel summation model.
[c] GCSM, general channel summation model.

distributions in going from one to two targets predict violations of Miller's inequality. This information is summarized in Table 3.

EXPERIMENTS

In conclusion of this section on capacity, we point out that it is logically possible for both the Miller and the Grice inequalities to be violated in the same data. However, such apparent contradictions can take place only at different intervals on the time axis. Furthermore, it is clear from the above developments that models classified by capacity cannot make predictions out of their domain (except perhaps by statistical artifact in low power data, etc.). For instance, an unlimited capacity independent channel parallel model cannot predict violation of either inequality anywhere; a channel summation model that preserves (or increases) its marginal distributions must predict a violation of Miller's inequality and is incapable of predicting violations of Grice's inequality and so on. Conversely, if an interval of response times is found to violate Miller's inequality in any model, the system must have been super capacity during that interval. Finally, it can be inferred from the above propositions that only a hybrid model in terms of capacity (and possibly architecture, etc.) could legitimately predict violations of both inequalities for the same data (but different time intervals as noted).

The two following experiments are nearly identical. For this reason, their method sections will be juxtaposed and the results considered together.

Two experiments, both employing the double factorial paradigm, with lighted dots as targets, were carried out. As noted above, this paradigm randomly intermixes trials where both targets are present, trials where the target is only on the right, trials where the target is only on the left, and trials with no target present. In the present realization, the first three trial types require a "yes" response and the last requires a "no" response. In addition, brightness of the dots was varied factorially. The double target trials, with varied brightness permitted analyses and models tests based on the systems-factorial technology developed above and in Appendix A. The comparison of single target and double target data allowed assessment of capacity relations that facilitate finer grain model discriminations (see Appendix B).

Each observer of each study participated in an experiment with nine possible display conditions: four brightness conditions for the redundant target condition, two brightness conditions for the left-only condition, two brightness conditions for the right-only condition, and one blank condition in which no stimulus was presented. All the experimental variables in both experiments were within-observer factors.

Method

Observers

In each experiment three (different) undergraduate students served as observers. In the first experiment, the observers were Purdue University students. In the second experiment the observers were students at Dartmouth College. All of the observers had normal or corrected-to-normal (20/20, with no known abnormalities, as reported by the observers) vision.

Apparatus

Experiment 1. A Scientific Prototype model N-1000/BA tachistoscope controlled the stimulus presentation. The luminance of the dot stimulus in the high brightness condition was $0.3\ \text{cd/m}^2$ and in the low brightness condition was $0.038\ \text{cd/m}^2$. The measurement of the brightness was accomplished by a Tektronix photometer. The dot size was 0.2°. The left dot was presented 0.5° to the left of the fixation point and 1.1° downward from the fixation point in the left eye, and the right dot was presented 0.5° to the right of the fixation point and 1.1° downward from the fixation point in the right eye. Hence, the dots stimulated non-corresponding points on the retinae. Separate channels of the tachistoscope were used for each eye so that each eye could receive either the left dot only or the right dot only. This resulted in dichoptic stimulation of the two eyes. No polarizing filter were employed in this experiment. No chin rest was necessary since the observers placed their forehead against the viewing goggles of the tachistoscope. An observer responded by pressing one of two response keys corresponding to "detection" response and "no detection" response.

Experiment 2. Three light emitting diodes (one green and two red LEDs) were installed on the optical bench which was placed 114 cm away from the viewer. The luminance of the target was controlled by a MacIntosh computer's A/D converter. The low luminance was $0.19\ \text{cd/m}^2$ and the high luminance was $9.98\ \text{cd/m}^2$. The sequence of the experiment as well as the data collection were controlled by the computer. One of the diodes served as a fixation point. The other two diodes served as target dots. The dot size was 0.5°. The target LEDs were placed in such a way that two lights were separated by 16.7° from each other. Thus, the two dots of light were again presented to non-corresponding points on the two retinae. In addition, two pairs of polarizing filters provided dichoptic stimulation of the two eyes. An observer responded by pressing one of two response keys, one corresponding to "detection" response and the other corresponding to "no detection" response. Response times were measured to an accuracy of 1 ms in both experiments.

Procedure

In Experiment 1, there were a total of 11 sessions: 7 practice and 4 data collecting sessions. A practice session consisted of 192 trials and a data collecting session consisted of 400 trials. Each session began with a 10-min dark adaptation period followed by 20 trials of practice. Each trial began with a presentation of a fixation point to each eye and a 1-s warning tone. The reaction timer started immediately after the end of this 1-s period and was stopped by the observer's response. On each trial there were four possible events: one dot in each eye; one in the left eye only; one in the right eye only; or none. The four events were equally probable. In the dot present conditions, the brightness of the dot was manipulated in such a way that the low brightness condition and the high brightness conditions were presented equally often. The observers were encouraged to respond as rapidly as possible while sufficiently evaluating the perceptual evidence as to permit a correct response.

In Experiment 2, four practice sessions were followed by four data collection sessions. Each session consisted of 480 trials with the target present and absent trials presented equally often. Within the target present condition the left-only, right-only, and redundant targets conditions were equally likely. Thus the probability of target absent $= \frac{1}{2}$, left only $= \frac{1}{6}$, right only $= \frac{1}{6}$, redundant $= \frac{1}{6}$. Each session began with a 5-min dark adaptation period. The parameters of stimulus presentation and response sequence were the same as that of the first experiment.

RESULTS AND ANALYSES

Initial Results: Mean RT ANOVA, Errors, and Level of Factorial Influence

The primary data of interest in the present experiments involve the redundant (double) target trials for brightness

TABLE 4

Mean RTs, Variances, and Error Percentages

	Both				Left		Right		
	ll	lh	hl	hh	l	h	l	h	No target
Mean RT	396	365	355	348	413	352	421	369	457
	422	376	363	357	433	363	441	376	564
	450	333	327	311	482	324	458	338	621
	393	261	276	246	436	283	442	264	602
	516	399	384	363	566	434	561	414	691
	510	393	413	369	615	456	598	424	639
Variance	827	995	545	675	1326	583	2070	1048	2240
	2364	2266	2803	1757	4189	2887	6354	2574	9292
	11328	7487	4798	3593	10487	5045	13480	7026	20912
	1935	910	1190	335	3963	1478	5001	681	7985
	16422	11221	7286	6724	9506	8365	12645	6139	62200
	14109	6615	8263	5077	18117	8827	14762	5476	19497
Error (%)	0.0	0.0	0.0	0.0	0.5	0.0	1.5	0.0	1.5
	0.0	0.0	0.0	0.0	0.0	0.5	1.0	0.0	1.5
	0.0	0.0	0.0	0.0	1.5	0.0	1.0	0.0	7.0
	0.0	1.3	0.0	0.0	3.9	1.9	1.9	0.6	1.2
	3.8	0.0	1.3	0.0	7.0	1.9	8.2	0.0	6.3
	5.0	0.0	0.0	0.0	9.4	1.9	6.3	0.0	1.7

Note. The data above each of the three lines pertains to the three observers in Experiment 1 (one for each row) whereas the data below each of the three lines pertains to the three observers of Experiment 2.

TABLE 5

ANOVA Table of Individual Observers

Source	DF	SS	F	P
Left intensity (L)	1	85030.6	122.85	0.0001
	1	152256.0	66.47	0.0001
	1	529911.2	79.91	0.0001
	1	352454.8	324.98	0.0000
	1	551013.7	57.23	0.0001
	1	28139.4	33.56	0.0001
Right intensity (R)	1	37210.4	53.76	0.0001
	1	68225.4	29.78	0.0001
	1	440033.2	66.36	0.0001
	1	506991.9	467.46	0.0000
	1	371925.7	38.63	0.0001
	1	507198.0	60.49	0.0001
L × R	3	13618.9	19.68	0.0001
	3	37791.4	16.5	0.0001
	3	256998.3	38.76	0.0001
	3	202204.1	186.44	0.0001
	3	180249.1	18.72	0.0001
	3	104749.3	12.49	0.0005
Day (D)	3	32343.1	15.58	0.0001
	3	23433.7	3.41	0.0177
	3	82387.1	4.14	0.006
	3	8849.6	2.72	0.0447
	3	279878.3	9.69	0.0001
	3	65232.9	2.59	0.0528
L × D	3	1969.5	0.95	0.4172
	3	1317.4	0.19	0.9020
	3	28347.5	1.42	0.2350
	3	6760.6	2.08	0.1032
	3	14002.4	0.48	0.6931
	3	5174.9	0.21	0.8924
R × D	3	349.1	0.17	0.9178
	3	1711.1	0.25	0.8620
	3	2405.6	0.17	0.9478
	3	815.7	0.25	0.8608
	3	45807.8	1.59	0.1928
	3	21690.7	0.86	0.4610
L × R × D	3	968.7	0.47	0.7058
	3	3933.1	0.57	0.6335
	3	34138.5	1.72	0.1632
	3	256.4	0.08	0.9714
	3	6976.8	0.24	0.8673
	3	27169.2	1.08	0.3578

Note. The results in the rows above the lines pertain to the three observers in Experiment 1 whereas those in the rows below the lines pertain to the three observers in Experiment 2.

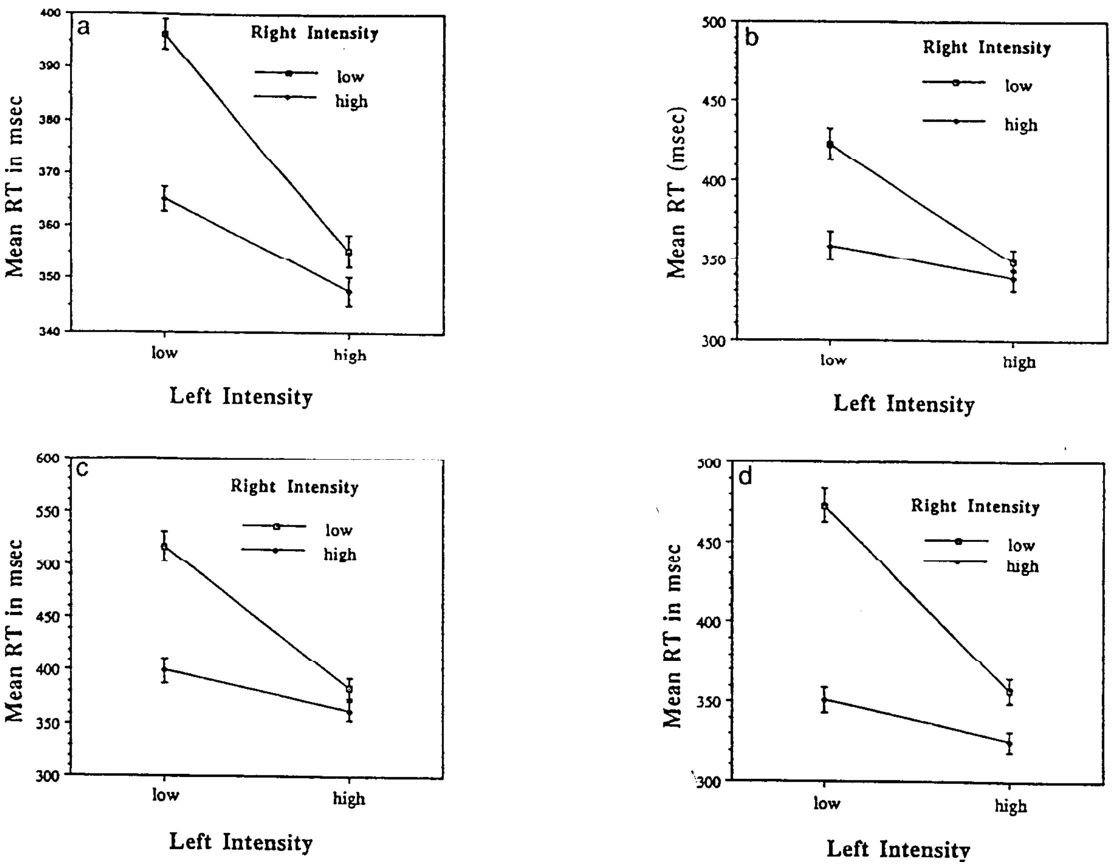

FIG. 4. Mean interaction contrast from Experiments 1 and 2: (a) A typical observer from Experiment 1. (b) The group average from Experiment 1. (c) A typical observer from Experiment 2. (d) The group average from Experiment 2. All are accompanied by standard error flags, with those of the group averages being the average standard errors rather than that from compiled data.

factor interaction analyses, and the one and two target trials for capacity analyses. The no-target trials are not of direct interest in this investigation.

All observers' mean RT, RT variance, and error rates in both Experiment 1 and Experiment 2 are listed in Table 4. As just noted, the ANOVA was confined to the 2×2 factorial subdesign contained within the set of trials on which both targets appear (left low or high × right low or high). Each observer's RT data for the redundant target trials were analyzed separately by ANOVA as shown in Table 5. The mean RTs of the redundant target condition for typical observers and averaged across observers are given in Fig. 4 for each experiment. The individual observers chosen for illustration in the figures were selected in an arbitrary fashion. The forms of the various functions were highly consistent across observers with minor exceptions to be noted and discussed.

The ANOVA results are very consistent across observers, all of which exhibited the following characteristics:

1. Main effects of the left and right conditions are highly significant (Ave. $p < 0.0001$ for both experiments).

2. The main effect of Days is significant (Ave. $p < 0.05$) at the 0.05 level, but, importantly, the interaction between Days and $L \times R$ is not (Ave. $p = 0.51$ for Experiment 1 and 0.73 for Experiment 2) nor are any of the other effects.

3. The mean interaction contrast, however, is positive: The $L \times R$ effect was significant at the $\alpha = 0.0001$ level for five observers, and $\alpha = 0.0005$ for one observer in Experiment 2. This is very strong evidence for mean RT overadditivity.

Thus, the initial evidence favors self-terminating (minimum processing time) parallelity or parallel channels pooled together downstream. It argues against seriality and exhaustive processing of either parallel or serial variety.

The percentage of errors for each observer is shown in Table 4. None of the observers in Experiment 1 responded incorrectly under the redundant target conditions and Experiment 2 observers made almost no errors under those conditions. However, in the single target conditions all the observers produced a small number of incorrect responses, ranging from 0 to 9.4% errors. In the single target condition, the low brightness condition yielded somewhat larger error proportions than the high brightness conditions. In the no-target condition, all observers evidenced relatively small 1 to 7% errors. Furthermore, the conditions under which the observer responded incorrectly are those under which the mean RTs are relatively slow. The no-target conditions are not of direct interest in the present study and will not undergo further analyses.

The level at which an experimental factor affects a distribution was broached earlier. Two relatively powerful levels are the density function one point crossover property and the ordering of survival functions. The former implies the latter but not the reverse. Figure 5b suggests that the stronger property of density one point crossover may be satisfied in addition to the weaker survivor function ordering. Of course, such data cannot guarantee that when both signals appear, such influences are totally selective, that is, they do not influence the inappropriate channel. But they do support a relatively strong level of channel influence.

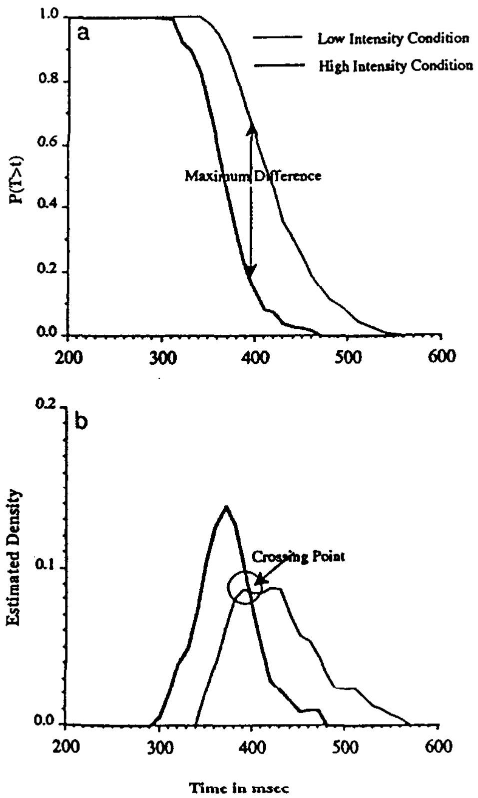

FIG. 5. Distributional influences of the high vs low intensity manipulation in a one-dot condition with (a) estimated survivor functions for a typical observer exhibiting survivor function ordering and (b) estimated densities for the same observer as in (a), suggesting satisfaction of the one point crossover. See text for further detail.

Tests of Survivor Function Interaction Predictions

For each observer, the estimated survivor functions of the four redundant target conditions were calculated simply as one minus the cumulative distribution functions. The latter are, of course, just the cumulated relative frequencies across time bins starting at the earliest point at which response times occur. Estimates of empirical cumulative distribution functions are standard and are not observed to have the difficulties associated with, say, estimates of density functions (e.g., Brunk, 1960, p. 131; cf. Silverman, 1986). We arbitrarily selected the reasonably intuitive (relative to the resolution of our models and the spread of the data) time bin interval of 10 ms.

Figure 6 shows the four survivor functions for the redundant dots case for the typical observers and grouped data in both experiments: [L(l), R(l)], [L(l), R(h)], [L(h), R(l)], [L(h), R(h)]. Observe that the ordering [L(h), R(h)] < [L(h), R(l)] ? [L(l), R(h)] < [L(l), R(l)]—where "?" indicates there is no predicted ordering by the models—holds well for the individual and group data, as required by selective influence at the survivor function level. Nonparametric Kolmogorov–Smirnov (see, e.g., Siegel, 1956; Kendall & Stuart, 1969) tests of predicted survivor function orderings, for example $[S_{R(l),L(l)}(t) - S_{R(l),L(h)}(t) > 0]$, were performed on each of the 12 pairs of distributions, 6 in each experiment, for each observer. In every case where an order was predicted, the appropriate ordering was supported (e.g., null hypothesis of $[S_{R(l),L(l)}(t) - S_{R(l),L(h)}(t) \leqslant 0]$ was statistically rejected) at the $\alpha = 0.01$ level.

Now abbreviate the survivor function interaction contrast as IC(t). Recall that IC(t) for the survivor functions is based on the same structure as the mean interaction contrast, except that it is applied to the survivor functions at each point (i.e., time bin) of time. Then the predictions by the serial and parallel models for the major stopping rules in the redundant target conditions were found to be: 1. Serial & self-terminating: $IC(t) = 0$ for all $t > 0$. 2. Serial & exhaustive: $IC(t) < 0$ for small t and $IC(t) > 0$ for large t. 3. Parallel & self-terminating: $IC(t) > 0$ for all $t > 0$. 4. Parallel & exhaustive: $IC(t) < 0$ for all $t > 0$. 5. Coactivation: $IC(t) < 0$ for small t and $IC(t) > 0$ for large t.

Figure 7 exhibits the survivor function interaction contrast for the same typical observers and the group data for the two experiments. Note that the interaction contrast functions are overwhelmingly positive. The exception is a negative excursion showing up in all but one of the graphs in Fig. 7. This result will be discussed further below. The massive positivity was found for every observer. No standard statistical tests are completely appropriate for assessing this statistic, but qualitatively the findings reveal massive negative evidence against serial self-termination, serial exhaustive, and parallel exhaustive processing.

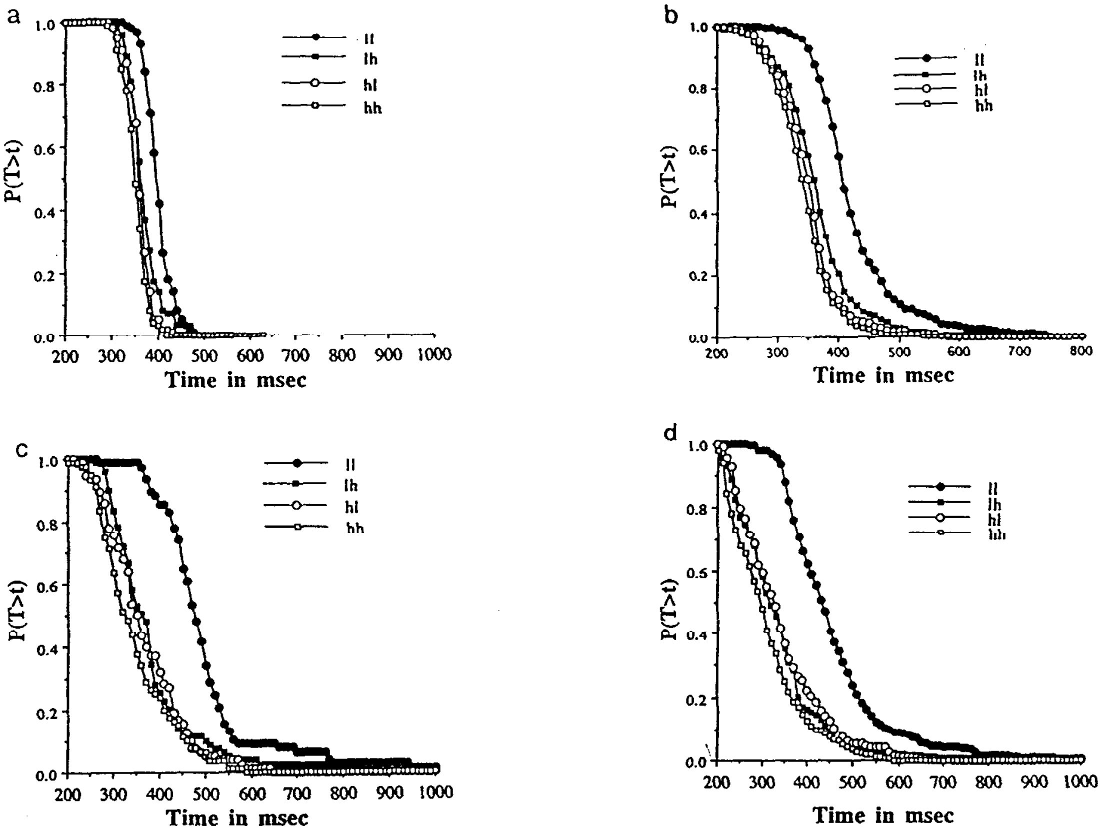

FIG. 6. Estimated survivor functions from the two experiments in the redundant targets (two-dot) conditions under the factorial intensity manipulations. (a) Typical observer from Experiment 1. (b) Group data from Experiment 1. (c) Typical observer from Experiment 1. (d) Group data from Experiment 2.

Standard parametric tests are compromised to some extent by the fact that successive RT frequencies from the estimated survivor functions are positively correlated. Attempts to compensate for these dependencies rapidly became extremely unwieldy. Nevertheless, the results of Z-tests are suggestive of the extreme positive trends in the survivor interaction contrasts. Details are given in Appendix C. The minimum overall Z score for an observer in either of the two experiments was $Z=3.95$ and the maximum was $Z=19.15$. Even with correlated survivor function frequency estimates there would be no a priori reason to expect the large-scale positivity of interaction contrast.

Nevertheless, the small negative blip at short RTs mentioned above occurred in five out of the six observers' data. The blips are invisible in the grouped Experiment 1 data because they naturally occurred at somewhat different points in time and Observer 3 in that experiment evidenced only positivity. Of all the models analyzed so far, only the exhaustive serial models and the coactivation channel summation models predict a negative deviation in the survivor interaction contrast function. And recall that exhaustive serial models were conclusively disconfirmed by the mean interaction contrast findings.

If processing were serial and exhaustive, for how small a t should we expect to find underadditivity? Assume that selective influence occurs at the density crossing level. That is, it is assumed that lowering the intensity moves a density function upwards (i.e., longer) in such a way that the densities cross only once. Let t_L and t_R be the density crossover points. That is, t_L is the point at which the high and low intensity density functions for the left side cross over, and similarly for t_R. Then, $\mathrm{IC}(t)$ should be negative in regions of time where t is less than both t_L and t_R. This can be easily seen in the expression and discussion of the independent (and order invariant) case following Proposition 3 and in the more general cases in Appendix A. Thus, negative interaction contrast should be found at t lying before the crossover t_L such that $f_{L(l)}(t_L)-f_{L(h)}(t_L)=0$ or before t_R such that $f_{R(l)}(t_R)-f_{R(h)}(t_R)=0$. If the left and right channels are independent, we can even estimate such points. Under the fairly natural additional assumption that the serial densities

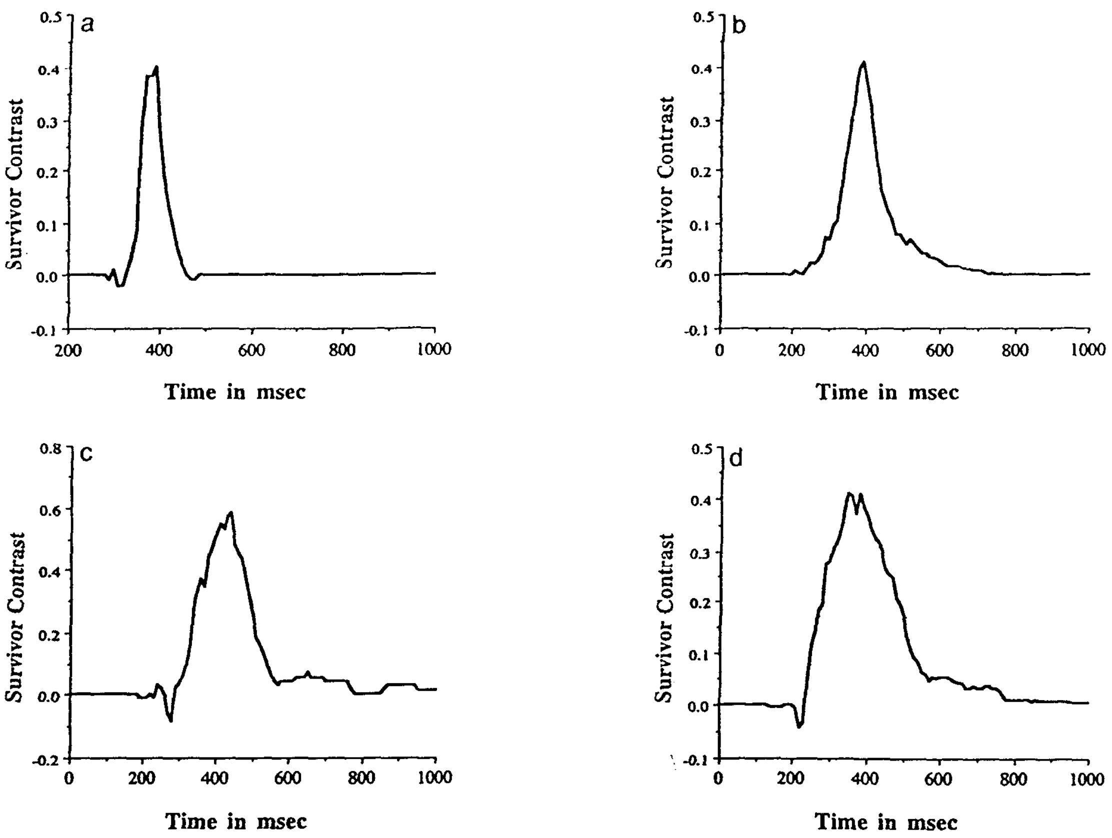

FIG. 7. The survivor function interaction contrasts are shown for typical observers and group data in the two experiments. See text for discussion.

are each unchanged in going from the one dot to redundant dot conditions, we can estimate the crossover points from the single dot conditions. Another helpful fact is that $S_{R(l)}(t) - S_{R(h)}(t)$ is maximal at $t = t_R$ and similarly for the left side L (see Fig. 5b).

Hence, t_R was estimated by locating the approximate maximum value of the difference $S_{R(l)}(t) - S_{R(h)}(t)$ and the crossing point for the left-only condition, t_L was estimated by locating the approximate maximum value of the differences $S_{L(l)}(t) - S_{L(h)}(t)$, in both instances, from the single dot conditions. For Experiment 1, the estimated values of t_L, t_R for Observer 1 were $\hat{t}_L = 380$ ms, $\hat{t}_R = 390$ ms, for the right, for Observer 2, $\hat{t}_L = \hat{t}_R = 390$ ms, and for Observer 3, $\hat{t}_L = 370$ ms, and $\hat{t}_R = 330$ ms. For Experiment 2 the values were $\hat{t}_L = \hat{t}_R = 310$ ms for Observer 1, $\hat{t}_L = 470$ ms, $\hat{t}_R = 450$ ms for Observer 2, and $\hat{t}_L = 500$ ms, $\hat{t}_R = 470$ ms for Observer 3. We next conducted a Z-test on $IC(t)$ for the $t \leq MAX(\hat{t}_L, \hat{t}_R)$ since it should be negative for all t less than either. (It was also performed for the more conservative $t \leq MIN(\hat{t}_L, \hat{t}_R)$, with the same results.) This test was performed on the average Z-scores for t obeying the above stipulations, for each observer in both experiments. The smallest Z was $Z = 7.07$ and the largest was $Z = 19.15$. Even with statistical bias due to correlated scores these values are notable and give no hint of negativity. Tests were also run previously at the crossover points where no bias would be present with sometimes smaller but still highly significant scores: MIN $Z = 3.2$, MAX $Z = 22.41$.

In contrast, the Poisson channel summation model predicts qualitative results much like those found in the five "negative-blip" observers, namely a small negative excursion far to the left within the overall survivor interaction contrast function. Figure 8 compares the behavior of a Poisson channel summation model with serial exhaustive, parallel self-terminating, parallel exhaustive and serial self-terminating processing models. The data of all five "negative blip" observers revealed maximum negative values qualitatively more in line with the Poisson channel summation model. Of course, the examples in Fig. 8 do not prove that all models in the various classes behave in "representative" manners. The theorems do.

Finally, the magnitudes of the negative deviations are massive in the exhaustive parallel and serial models as follows from the theory and as can be seen in Fig. 8 (refer to

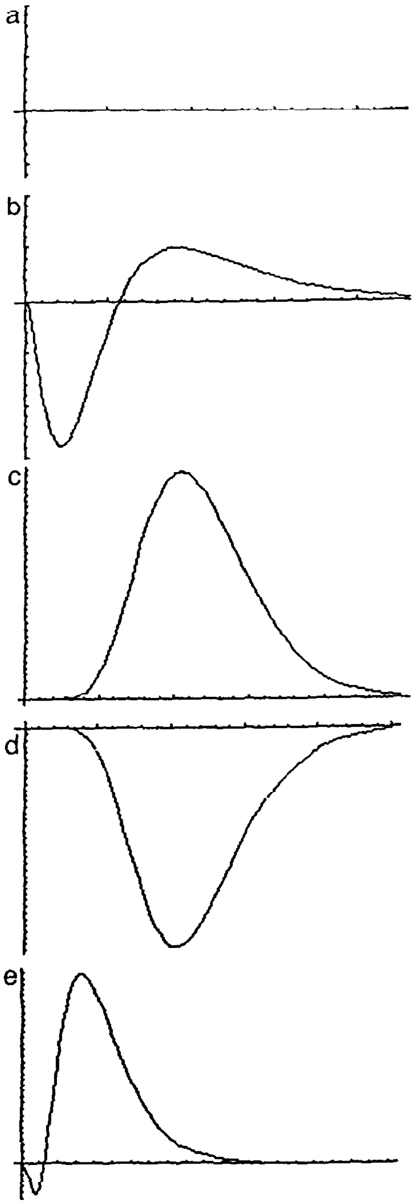

FIG. 8. Representative predictions of survivor interaction contrast by (a) a serial self-terminating (minimum time) model, (b) a serial exhaustive model, (c) a parallel self-terminating (minimum time) model, (d) a parallel exhaustive model, and (e) a channel summation (coactive) model. All the illustrative models are based on Poisson counting distributions.

the discussion after Proposition 5). They are much smaller in the Poisson summation models, relative to the positive region, as they are in the data. In fact, the negative regions for the channel summation models can never equal the positive area, as must the serial exhaustive models (see Propositions 4 and 5 and their discussion), which in any case were decisively falsified by the mean interaction contrast data. Hence, the interaction contrast data are clearly closest to the coactivational channel summation predictions.

In summary, investigation of survivor function interaction contrast is most supportive of parallel channels either with self-termination or more likely with a summation of channel information indicating coactivation.

Assessment of Redundant Targets and Capacity Effects

As submitted earlier, we view the effects of manipulating the number of redundant targets on RT as reflecting capacity properties of the system. In exhaustive parallel processing, statistical effects that act to prolong waiting for the last item to complete basically augment any limitations in individual channel capacity, both serving to elongate RTs (e.g., Townsend, 1974). In the present experimental design, the first target item finished can determine the response. However, the statistical (horse race) advantage associated with the minimum completion time random variable can, in principle, be offset by the increase in perceptual (cognitive, etc.) load in moving from one to two presented items, as discussed earlier. Hence, the overall capacity of the system, involving a tradeoff of statistical (increase efficiency), load effects (decrease efficiency), and perhaps coactivation (increase efficiency), determine the outcome, and the resultant is what is being "measured" by $C(t)$, Miller's and Grice's inequalities.

In assessing the difference in performance, in going from one dot to two dots, we employ the inequalities presented earlier as well as estimates of capacity from the survivor function data. First we compare the performance in the two-target conditions with that expected on the basis of independent unlimited capacity parallel processing. Figures 9a and 9b estimate $\hat{C}(t)$ for the two experiments. Note that Experiment 1 appears to support a degree of limited capacity except for a few small values of t. That is, $\hat{C}(t) < 1$. The geometric means of $C(t)$ across time bins varied between 0.575 and 1.02 across individual observers and conditions with all but two being less than 1.00. Recall that fixed capacity is defined by the average of the two single target integrated hazard functions, which implies the spreading of a fixed amount of processing capacity over the two positions. Equivalently, fixed capacity means that $C(t) = \frac{1}{2}$, with smaller values of $C(t)$ implying extremely limited capacity. And although the geometric means of $C(t)$ do not fall below fixed capacity ($C(t) = \frac{1}{2}$), many values hover around that value. On the other hand, Experiment 2 appears to signify considerable super capacity, $\hat{C}(t) > 1$, especially at the smaller RTs. The geometric means of $C(t)$ across time bins ranged from 0.730 to 1.926, with only 2 out of 15 being less than 1.

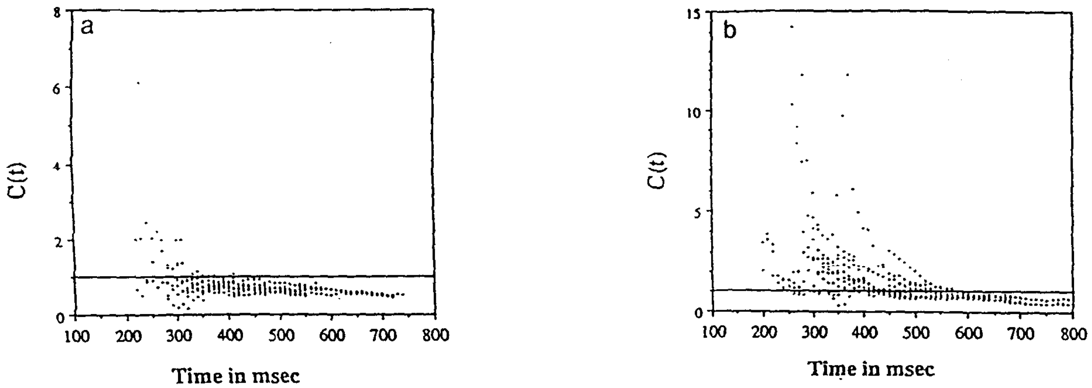

FIG. 9. Capacity coefficient plots for (a) all observers and times in Experiment 1 and (b) all observers and times in Experiment 2. See text for details.

The $C(t)$ measure has told us that capacity was limited in Experiment 1 and super in Experiment 2. However, Grice's inequality provides a bound that, if exceeded, indicates very limited capacity whereas Miller's inequality provides a bound that, if exceeded, indicates a highly super capacity system in operation.

With regard to the Grice inequality, recall that if the survivor function in the redundant target (i.e., two dot) condition, $S_{\mathrm{L,R}}(t)$ satisfies $\mathrm{MIN}[S_{\mathrm{L}}(t), S_{\mathrm{R}}(t)] - S_{\mathrm{L,R}}(t) < 0$, then the capacity function

$$C(t) < \frac{\mathrm{MAX}[H_{\mathrm{L}}(t), H_{\mathrm{R}}(t)]}{H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t)}, \quad \text{where } H_i(t)\ (i = \mathrm{L, R})$$

are the respective integrated hazard functions from the single target conditions. This implies moderately limited to extremely limited capacity.

Figures 10a and 10b give the results for the two experiments. The area of interest here is that lying below 0. Observe that Experiment 1 yields evidence for consistent capacity limitations more severe than that imposed by the bound $\mathrm{MIN}[S_{\mathrm{L}}(t), S_{\mathrm{R}}(t)]$. On the other hand, there are relatively few violations of this inequality by the data of Experiment 2. We will see further evidence compatible with Figs. 9 and 10 in the next analysis.

Miller's inequality is examined in Figs. 11a and 11b with the computation of $S_{\mathrm{L,R}}(t) - S_{\mathrm{L}}(t) - S_{\mathrm{R}}(t) + 1$ with the overwhelmingly observed positive values arguing strongly against unlimited or supercapacity in Experiment 1 and just the opposite occurring in Experiment 2. Thus, we conclude that in Experiment 1, $C(t) < 1$ and in some cases, $C(t) < \mathrm{MAX}[H_{\mathrm{L}}(t), H_{\mathrm{R}}(t)]/(H_{\mathrm{L}}(t) + H_{\mathrm{R}}(t))$ arguing for mild to very limited capacity in that experiment, across the range of response times.

In contrast, in Experiment 2 we find that the bound associated with the Miller inequality was violated numerous times, suggesting a strongly super capacity system.

There are two final points worth noting here. The present approach to redundant targets designs, couched explicitly in terms of capacity, has not been the usual way of thinking about the topic. We believe that the mathematics and logic, when used to embed the psychological concepts as we have done, justify this conceptual tactic. Further, the quantitative demonstration of the close linkage of such capacity indicants as the integrated hazard function, the coefficient $C(t)$, and the two major inequalities also support the thesis that capacity effects, both statistical and substantive, are the pivotal concepts here.

The propositions suggest that Miller's inequality at a time point t is not equivalent to super capacity. The inequality implies super capacity, but super capacity (in terms of $C(t)$) seems to be needed over a range of times (in Proposition 6, the full range) to imply a violation of the inequality. The fact that the various inequalities and statistics, coming from different angles of attack, all lead to the same basic implications should provide a strong measure of confidence in the conclusions.

Technical Considerations

The above calculations ignore the contributions from the "nonprocessing" components of the base or "residual" duration to RT. Consider the formula for $C(t)$ that would be appropriate, even with an independent base time included, if we were first able to deconvolve it from the left and right single target conditions. The derivation is analogous to similar developments in the appendices.

$$C(t) = \frac{-\ln\left[S_{\mathrm{B}}(t) + \int_0^t f_{\mathrm{B}}(t - t')\, S_{\mathrm{L,R}}(t')\, dt'\right]}{-\ln\left[S_{\mathrm{B}}(t) + \int_0^t f_{\mathrm{B}}(t - t')\, S_{\mathrm{L}}(t')\, S_{\mathrm{R}}(t')\, dt'\right]}.$$

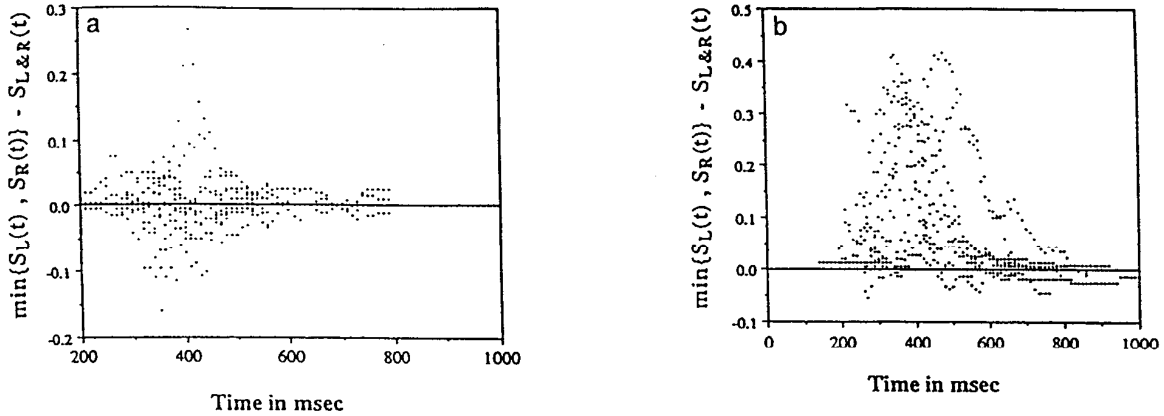

FIG. 10. Plots that exhibit potential violations of the Grice inequality for (a) Experiment 1 and (b) Experiment 2. See text for details.

However, there are technical difficulties in accurately assaying the contribution of the base time, even when it is independent of the processing times. Much progress has been made on this problem recently (e.g., Burbeck & Luce, 1982; Smith, 1990). There are still difficulties that attend this enterprise, as Smith's (1990) discussion makes apparent. Also, we shall see that some new theoretical work permits us to draw some powerful conclusions without that information, although this does not argue against the use of finer grain tests, when such are appropriate.

The actual formula that represents our estimates of $C(t)$ includes the base time in both the left- and right-hand distributions so that in our analyses,

$$C(t) = \frac{-\ln[S_B(t) + \int_0^t f_B(t-t')\, S_{L,R}(t')\, dt']}{\begin{bmatrix} -\ln[S_B(t) + \int_0^t f_B(t-t')\, S_L(t')\, dt'] \\ -\ln[S_B(t) + \int_0^t f_B(t-t')\, S_R(t')\, dt'] \end{bmatrix}}.$$

Hence, the base time appears twice in the denominator of $C(t)$. Now, it is straightforward to show, using Eq. (24) of Colonius (1990b; see also 1986), that if capacity is limited such that $C(t) < 1$ when the formula is devoid of the base time, then $C(t)$ will still be less than one with our less accurate measure. Furthermore, computer simulations suggest that our $C(t)$ will generally be less than that predicted by the distributions without the base time influence. So far, we have not been able to prove a general theorem to that effect and hence are not in a position to claim that diminution on $C(t)$ occurs with certainty.

Fortunately, it turns out that our major qualitative conclusions, drawn as they are from comparisons of the redundant target condition distribution with the Miller and Grice bounds on survivor functions, are impervious to contamination by statistical dependence between processing channels as well as to that by an independent base time. First, Colonius (1990b) has shown that if the individual marginal distributions are unchanged in the double versus single target conditions, statistical dependence cannot make performance exceed those bounds. Put another way, correlated parallel channels cannot cause violations of the Miller inequality, unless capacity is altered so greatly that the marginal distributions are changed as well. This fact is compatible with Proposition 6, revealing the intimate relation between violation of Miller's inequality and consistent or strong super capacity.

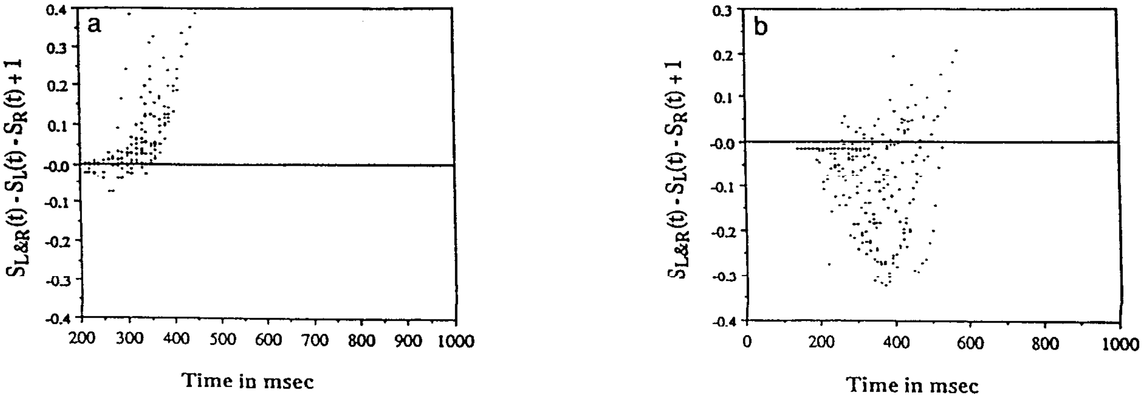

FIG. 11. Plots that exhibit potential violations of the Miller inequality for (a) Experiment 1 and (b) Experiment 2. See text for details.

Furthermore, Ulrich & Giray (1986) have demonstrated that the Miller and Grice bounds can be violated by a redundant target distribution including the base time, if and only if they are violated in their pure form (i.e., without being convolved with the base time distribution). Putting the Colonius (1990b) results together with those of Ulrich & Giray (1986) allows our major conclusions to stand firm: Experiment 1 produced capacity ranging from close to or a little less than fixed to something less than coactivation-level supercapacity. Experiment 2 evidenced super capacity in excess of that required for violation of Miller's inequality and hence is compatible with supercapacity parallel processing or coactive channel summation.

We end this section with a very brief discussion of variance as a function of a number of items. The theory pertaining to variance is not as well developed as that for mean RTs. Mean RTs are quite well understood with regard to parallel versus serial processing (e.g., Townsend & Ashby, 1983) and self-terminating versus exhaustive processing (Townsend & Van Zandt, 1990; Van Zandt & Townsend, 1993). Although the classes of models for which behavior of variances are limited, there are suggestive hints (e.g. Schneider & Shiffrin, 1977; Townsend & Ashby, 1983; Krueger, 1984). For instance, it is obvious that *standard* serial models predict that the variance for a minimum time stopping rule (the first item, a target, processed always stops processing) would remain constant across changes in the number of items (targets). When target rates differ over position (a non-standard serial model), there is predicted to be a slight-to-moderate increase in variance over a weighted average of the two individual variances. Obviously, exhaustive variance increases dramatically for such models, both standard and non-standard.

Parallel models have not been so well studied but related literature suggests that variance predictions can depend on the exact form of the individual item distributions (e.g., Gumbel, 1958). It *is* known that symmetric distributions produce an increase in variance in minimum time processing. On the other hand, at least certain common distributions, such as the exponential, predict an increase in variance as n grows with exhaustive processing and a *decrease* with minimum time (i.e., self-terminating on redundant target trials) processing. The present variances, as can be seen in Table 1, are closer to what would be predicted by the latter type of minimum-time parallel (race, e.g., exponential) models than exhaustive parallel or any usual type of serial models. Nevertheless, considerably more theoretical and empirical research on this topic is called for.

GENERAL SUMMARY AND DISCUSSION

Areas of Application and Binocular–Dichoptic Detection

The theory and associated methodology developed in this investigation can be applied in any situation where the major issues concern serial versus parallel processing and the form of the stopping rule, when the information sufficient to allow a correct response is available after the first item is completed. In addition, the method also permits assessment of the coactivation hypothesis. There are, however, certain research areas which have a history of interest in such problems and for which the present theory may provide a considerably enriched technology.

One such area is binocular detection, in particular, the question of whether improvement in performance found when the information is sent over two channels can be explained entirely by probability summation, that is, that improvement predicted by unlimited capacity, independent parallel models with nonzero variance of processing times. A detailed treatment of this specific research domain cannot be given here, but typically, monocular and binocular (sometimes dichoptic) detection of interaction contrast increments of light flashes (sometimes in the form of dots) or sinusoidal gratings are studied using RT, as in the present case, or accuracy methods (e.g., Blake, Martens, Garrett, & Westendorf, 1980; Legge, 1984; Westendorf & Blake, 1988). The double factorial paradigm and the associated analyses had not heretofore been employed in this domain and the capacity function was not previously available.

The results of the one versus two signal conditions have appeared to depend on whether the two eyes were binocularly (not dichoptically) stimulated with identical patterns on corresponding retinal areas. If so, the typical result was that binocular performance exceeded that expected on the basis of probability summation. Of course, whether or not Miller's inequality was surmounted was not addressed in the earlier literature. Contrarily, if the two eyes were stimulated at non-corresponding retinal areas or were stimulated with nonidentical patterns, binocular performance was found to be equivalent to that expected merely on the basis of probability summation (e.g., see Blake, Martens, & Di Gianfilippo, 1980). Both kinds of results tend to contradict seriality and support self-termination, but without the added support of the double factorial paradigm.

In the present case, the factorial results as well as the capacity findings support self-terminating parallelity within a range of mildly limited to very super capacity (i.e., super capacity sufficient to recommend coactivation). The capacity results differed across the two experiments, with Experiment 1 producing quite limited capacity versus the substantial super capacity found in Experiment 2. Could there have been some "leakage" through the polarizing filters in Experiment 2 to the unstimulated eye? Actually, such leakage might have been expected to lessen the appearance of super capacity: Leakage might be expected to aid detection more in the monocular case than in the dichoptic case where both eyes are already receiving high-energy signals. Further, during informal pilot tests and training, neither observers nor experimenters could detect any energy in the unstimulated eye.

In contrast, it may be that the moderate to severe limitations in capacity found in Experiment 1, $C(t) < 1$ and violations of Grice's inequality, were due to contamination introduced by inclusion of the base times in estimates of the capacity function $C(t)$ or a positive dependence (although we view the latter as unlikely) as discussed in the previous section. Hence, in Experiment 1, it may be that some of the limitations in capacity were perpetrated by the base time contamination, but the rather frequent incursions of data below Grice's bound suggest that some of the capacity degradation was due to load effects and possibly some positive correlation. Indeed, it is not impossible, as observed in the theoretical sections, that coactivation together with extremely limited capacity could have been responsible for Experiment 1 results.

With regard to the binocular detection literature, the present dichoptic, non-corresponding retinal stimulation might lead one to expect probability summation and neither limited nor super capacity. In addition, the dichoptic nature of the stimuli argues against a peripheral "neural summation" interpretation of coactivation. This type of result is one reason we generally favor the neutral term "channel summation" to represent the general class of models that includes the Schwarz (1989) superposition model. Nevertheless, there could be physiological interaction at higher neural levels. The precise psychophysical or psychophysiological causes of present results including the differences between Experiments 1 and 2 are not the primary goal of the present study but are undergoing continued investigation. More discussion of the latter topic appears below.

Other current research areas beginning to apply mathematical models of parallel and coactive processing, and which have close ties to psychophysics and physiology, include the study of bimodal sensory integration (e.g., Diederich, 1992; Hughes, Reuter-Lorenz, & Nozawa, 1993; Nozawa *et al.*, 1993).

Other Related Studies

The literature on redundant signal effects is voluminous and the results one can expect likely depend on whether the stimuli are presented within the same modality, the nature of the stimulus materials and other contextual and perhaps observer factors. This is true even within a relatively homogeneous milieu, such as that involving experimentation using elementary visual symbols. Whether the assessed capacity is unlimited capacity parallel or limited capacity (serial or parallel) varies with the stimulus patterns and task characteristics (e.g., Egeth *et al.*, 1988; Fournier & Eriksen, 1990; van der Heijden *et al.*, 1983). The purpose here has not been to review the literature but to focus on the development of a coherent theory and associated methodology along with an intensive exploration of that technology within a straightforward experimental setting. Nevertheless, it is important to relate our study to recent investigations of a generally similar nature. We start with a brief survey of other mathematical models for redundant target situations. None of these were developed in terms of factorial predictions.

The more mathematized models applied to previous data have been: 1. Grice and colleagues' parallel deterministic growth—with random decision threshold model (Grice *et al.*). 2. The Miller (1986) coactivation and its hazard function specification by Colonius (1988). 3. The superposition model of Schwarz (1989), further investigated by Diederich & Colonius (1991) and Colonius (1990a). 4. The several models tested by Meijers & Eijkman (1977).

The superposition model is a special case of the class of summation models developed here. It is equivalent to the Poisson channel summation model. Both the Grice and Schwarz models have difficulty with some aspects of various data, but handle other aspects quite well. For instance, Diederich & Colonius (1991) and Colonius (1990a) have continued to test the model of Schwarz within bimodal and binocular contexts and thus further refine understanding in these areas. The Miller (1986) coactivation model does not pin down the nature of the redundant target hazard function and thus can predict any type of capacity result. The Meijers & Eijkman (1977) study favored a special type of parallel independent race model.

Next, Fournier & Eriksen (1990) review several reasons why coactivation operationally defined in terms of Miller's inequality might not appear. The first reason is that of limited capacity which reduces speed on individual items or channels when two rather than one are present. This is exactly the conclusion to be drawn if other explanations can be ruled out. But it leaves unexplained why this would occur in Experiment 1 but not in 2.

The second reason suggested by Fournier & Eriksen (1990) is the possibility of a correlation "between the parallel processing of each target on redundant presentation due to a common third variable." This point may be related to the stochastic dependency effects discussed at length below in regard to studies by Mordkoff & Yantis (1991) and Miller (1991) and in earlier sections, the major difference being that the Fournier and Eriksen correlation derives from a third influence rather than being inherent in the processes.

Fast guessing was suggested as a third factor that might decrease performance in the redundant target conditions. Given the high level of performance in the double dot conditions, this does not seem a likely explanation. It also does not encompass the superior performance in Experiment 2. The Fournier & Eriksen (1990) work, in alliance with other studies, also points up the importance of stimulus complexity. This factor is minimal in the present experiments.

With regard to the possible coactivation found in Experiment 1, Fournier and Eriksen observe (along with other writers) that serial processing could lead to faster RTs on the double target condition, because on single dot trials, half the time the wrong position would be processed first. This explanation is doubtful here by virtue of the factorial results, which found extremely strong evidence for overadditivity and not additivity; the latter being implied by a very large class of serial models.

It is germane to mention that Dzharafov (1992) has recently developed a class of models that permit strong dependencies between the base time and the processing durations. Although that paper was primarily confined to simple RT, such assumptions in the context of parallel race might complicate or actually simplify analyses such as those carried out here, depending on the specific versions employed. Dzhafarov (1993) investigates general parallel race models, especially with regard to deterministic criteria plus stochastic activation versus stochastic criteria plus deterministic activation. Colonius & Townsend (1992) provide a general mathematical characterization of race versus coactivation models and discuss the implications of Dzharafov's (1993) results and other aspects of testability.

Egeth & Dagenbach (1991) carried out a double factorial study using a design closely related to that here and those in our own laboratory (Nozawa, 1989; Townsend & Nozawa, 1988; Townsend & Piotrowski, 1981). An interesting difference was that it employed one type of signal as target (e.g., X) and another as non-target or distractor (e.g., 0), rather than a target present (i.e., dot) versus target absent (i.e., blank). Thus, if either position contained an X, a "yes" response was made whereas a "no" response was made only if both positions contained an 0. This feature permits two tests of parallel models: overadditivity on redundant target trials (e.g., two X's present) where minimum time processing could occur, as well as underadditivity on redundant nontarget trials (e.g., two 0's present) which forced exhaustive processing of the nontargets. Indeed, the investigator could also test for additivity on single target trials (e.g., an X and an 0 present), but this prediction is made by both self-terminating serial and self-terminating parallel models, so is somewhat less interesting as a diagnostic.

We focus on their first two experiments. Egeth & Dagenbach (1991) found that the underadditivity appeared as predicted in an experimental setting designed to encourage parallelity but not in one where serial processing might be expected. Unlike the present experiments, the overadditivity diagnostic was less helpful in their context. It is also important to notice that since interaction contrast was only examined at the mean level and capacity effects were not analyzed, there was no diagnostic available to discriminate coactivation from parallel race models.

Another recent paper with important implications for redundant target experiments is that by Mordkoff & Yantis (1991). Although they do not employ the double factorial paradigm, they advance a qualitative model termed the "interactive race model" which permits information transfer at two different points between parallel channels. This structure permits one channel to adjust its processing parameters within trials according to information it is receiving about the signal on the alternate channel. Their several experiments decisively show that RTs can be affected by experimenter-determined signal dependencies, sometimes leading to violations of Miller's inequality. The authors point out that the interactive race model can produce race gains that violate Miller's inequality. This becomes apparent in the quantitative treatment by Colonius & Townsend (1992), where it is shown that structurally speaking, coactivation models are a special case of interactive race models. The converse is not true within the definitional apparatus advanced here. That is, parallel race models, as conceived herein, are not a mathematical special case of coactive models. We thus depart from Miller (1991, p. 167) on this issue.

Another aspect of the Mordkoff and Yantis model is their assumption that in the absence of certain stimulus–stimulus or stimulus–response contingencies, the predictions must obey Miller's inequality. More discussion on this will follow later.

Mordkoff & Yantis (1991) find, in surveying the studies where Miller's inequality had been violated, that all had inadvertently (usually) included intersignal or signal-response dependencies (contingencies) that could have led to interchannel interactions of the sort encompassed by their model. For present circumstances, the critical aspect of the Mordkoff and Yantis argument occurs when the conditional probability of, say, a target on the left given a *non*target on the right, is less than the marginal probability of a target on the left. The reasoning is that when the right nontarget happens to be completed first, the information conveyed to the left channel is that a target is less probable than average, and a slowdown occurs. On the other hand, the very first item completed on a redundant target trial immediately determines the "yes" response and no slowdown occurs. Thus, the comparison of redundant versus single target trials might reveal the spurious appearance of a "super capacity" effect.

In our Experiment 1, the stimulus probabilities were such that no information could be conveyed by a nontarget (blank) in either location, and no supercapacity result was found then. By contrast, the conditional probability of a target on the left given a nontarget on the right was 0.25 compared with the marginal probability of a target on the left of 0.33. Thus, the conditions in Experiment 2 were right for a mechanism of the sort envisioned by Mordkoff & Yantis (1991). We shall discuss this possibility below.

A key provision for testing the interactive race model against the coactivation model has been the postulate by

Mordkoff & Yantis (1991) that the channels of the race model interact only if interstimulus contingencies favor the presence or absence of a target or of a response over another (e.g., a nonresponse). Thus, in the absence of such contingencies, the interactions are in effect shut down. This assumption is not entailed by the parallel structure or the possibility of interactive channels. For instance, two parallel channels could sum their information within each channel, perhaps with information (e.g., counts) weighted less that is siphoned from the "other" channel. In fact, as noted earlier, within the context of more general interactive parallel models, Colonius & Townsend (1992) show that a general interactive parallel race model can easily degenerate into a coactivation model, by permitting cross-talk but closing off one of the ensuant parallel channels, thus enabling it to mimic the latter's interaction contrast and capacity predictions. It appears to the present authors that the assumption that without interstimulus or stimulus-response contingencies present, no "coactivation" occurs, is not an inherent aspect of interactive parallel models.

Miller (1991) has recently proposed a model that is the coactive analogue to the Mordkoff & Yantis (1991) approach. That is, two parallel channels might interact before summing their input. He also discusses the possibility that the two components of a stimulus configuration are consolidated into a unified Gestalt-like pattern. Miller's experiments go on to offer evidence that Gestalt-like effects can occur across certain stimulus dimensions that are incompatible with the notion of independent channels. Of course, such evidence is not necessarily only associated with coactivation, but possibly with interactive parallelism, although not in the version put forth by Mordkoff & Yantis (1991).

Summing up the most recent evidence (emphasizing aspects of particular relevance to the present study): Mordkoff & Egeth (in press) find that several studies involving RT or accuracy offer support for the standard interactive race model. Mordkoff & Miller (1993) show that Miller's inequality may be violated even though no biased contingencies are present. This falsifies the standard Mordkoff & Yantis (1991) race model but not more general parallel models and certainly not coactive models.

Nozawa (1992) has demonstrated in a binocular double factorial paradigm, employing both factorial and capacity analyses, evidence for limited capacity parallelity or coactivation (that is, combination of channel information, but with less capacity when both channels are utilized) in the redundant target part of the paradigm. Evidence for exhaustive parallel processing appeared in the double nontarget trial data (both positions being occupied by nontarget information requires exhaustive processing for high accuracy). In an accuracy version of the design, support was garnered for unlimited capacity, parallel, probability summation under the redundant target condition and a super-additive. multiplicative type of coactivation under the no-target (hence, exhaustive processing needed for perfect accuracy) condition.

Whatever the characteristics of interactive race models or interactive coactivation models, it is important to reemphasize certain facts. First, any such model cannot violate Miller's inequality and still maintain marginal RTs that are equal to those observed on single signal trials. That is, the stochastic dependencies themselves cannot produce such extreme effects. This result was proven by Colonius (1990b; see Proposition 4.8, p. 266 and also Colonius, 1986): Such models have a bound which is precisely the edge of Miller's inequality, as well as another bound identical to the Grice inequality. Hence, to predict Miller's inequality violations, the information transfer has to actually affect capacity by speeding up the individual channels, in the sense of altering the marginal processing distribution on the channels. Because interactive parallel models which leave the marginal distributions unchanged obey the Miller and Grice inequalities, perhaps they deserve a special name, such as "conservative interaction parallel (race) models."

We briefly review the results of our analyses, followed by conclusions. 1. The mean interaction contrast results support self-terminating parallel or channel summation processing and strongly disconfirm exhaustive serial and exhaustive parallel processing. 2. The survivor function interaction contrast results strengthen the mean interaction contrast findings and by parsimony and present knowledge of model behavior are most supportive of channel summation processing. The positive interaction contrast for very small, moderate, and large RTs, accompanied by the small negative blip at moderately short times, are predicted by the Poisson channel summation model (i.e., the contemporary superposition model) and a class of other channel summation models. The exhaustive serial class of models incorrectly predicts very large negative survivor interaction contrast (equal in area to the positive interaction contrast) for a wide range of smaller RTs, and they cannot predict the interaction contrast results. The self-terminating parallel models covered here are unable to predict any negativity in interaction contrast. 3. The moderately to severely limited capacity effects in Experiment 1 and the super capacity effects in Experiment 2 both could have been produced either by a self-terminating interactive parallel race or by a mildly limited capacity, parallel channel summation model. The channel summation model may be a bit more parsimonious here.

In conclusion, our results strongly affirm parallel nonexhaustive processing. The balance of evidence favors an interpretation where the parallel channel information is summed (coactivation) as opposed to self-terminating (here implying minimum time processing) parallel race models, particularly within the mathematically specified models developed herein. Effective capacity at the level of moderately to very limited was found in Experiment 1 and super capacity in Experiment 2 was supported. Serial and all exhaustive processing models were decisively falsified. Although several plausible reasons for the disparity in capacity measurement in the two experiments were discussed, a conclusive determination remains to be found. Considerably more research will likely be required to decisively test the most general interactive race and coactive models, as well as to establish the site and nature of impact of the important experimental parameters being unearthed.

APPENDIX A

On the Spatio-temporal Properties of Elementary Perception: An Investigation of Parallel, Serial, and Coactive Theories

Parallel Process Models

As in the text, we designate joint events or values by a comma. Thus, $f(t_{\mathrm{L}}, t_{\mathrm{R}})$ indicates a bivariate density with values at t_{L} and t_{R}, and $P(T_{\mathrm{L}} \leqslant t_{\mathrm{L}}, T_{\mathrm{R}} \leqslant t_{\mathrm{R}})$ illustrates a joint cumulative distribution function. An "OR" event will simply be designated by the term "or" within an expression.

Suppose that the marginal base time distribution function $f_{\mathrm{B}}(t)$ is invariant across factor levels of L, R. Next consider the joint density for L, R based on factor levels *i*, *j*, respectively, and conditional on the base time t_{B}

$$\begin{aligned}&f_{\mathrm{L}(i),\,\mathrm{R}(j)}(t_{\mathrm{L}}, t_{\mathrm{R}} \mid t_{\mathrm{B}})\\ &\quad = f_{\mathrm{R}(j),\,i}(t_{\mathrm{R}} \mid t_{\mathrm{B}})\, f_{\mathrm{L}(i),\,j}(t_{\mathrm{L}} \mid \mathrm{T}_{\mathrm{R}(j)} = t_{\mathrm{R}}, t_{\mathrm{B}}),\end{aligned}$$

where the second subscripts in the expressions on the right indicate a potential direct dependence of the respective functions on the level of the other factor. Thus, $f_{\mathrm{R}(j),\,i}(t_{\mathrm{R}})$ is the marginal density for the right side (itself at level *j*) in the presence of level *i* on the left. We now assume that $f_{\mathrm{R}(j),\,\mathrm{l}}(t_{\mathrm{R}} \mid t_{\mathrm{B}}) = f_{\mathrm{R}(j),\,\mathrm{h}}(t_{\mathrm{R}} \mid t_{\mathrm{B}})$, that is, the R density integrated over t_{L}, but conditioned on t_{B} is invariant over the two L factor levels, l, h. Moreover, we assume that

$$\begin{aligned}&f_{\mathrm{L}(i),\,\mathrm{l}}(t_{\mathrm{L}} \mid \mathrm{T}_{\mathrm{R}(\mathrm{l})} = t_{\mathrm{R}}, t_{\mathrm{B}})\\ &\quad = f_{\mathrm{L}(i),\,\mathrm{h}}(t_{\mathrm{L}} \mid \mathrm{T}_{\mathrm{R}(\mathrm{h})} = t_{\mathrm{R}}, t_{\mathrm{B}})\\ &\quad = f_{\mathrm{L}(\mathrm{i})}(t_{\mathrm{L}} \mid \mathrm{T}_{\mathrm{R}} = t_{\mathrm{R}}, t_{\mathrm{B}}).\end{aligned}$$

That is, the left conditional density depends only on the actual processing time of the right side, not explicitly on the factorial level of R. Of course, it can also depend on t_{B} as the notation indicates. Note too that the posited invariance of densities implies invariance at the survivor level.

It is assumed finally that selective influence operates at the conditional survivor ordering level. That is,

$$S_{\mathrm{R}(\mathrm{l})}(t_{\mathrm{R}} \mid t_{\mathrm{L}}, t_{\mathrm{B}}) > S_{\mathrm{R}(\mathrm{h})}(t_{\mathrm{R}} \mid t_{\mathrm{L}}, t_{\mathrm{B}}) \qquad \text{for} \quad t_{\mathrm{L}}, t_{\mathrm{R}}, t_{\mathrm{B}} > 0.$$

Some of the theorems will depend on a stronger level of selective influence. Let $\Delta^2_{\mathrm{L},\,\mathrm{R}}$ represent the second-order mixed partial difference over the factor levels, or contrast operator, which corresponds to the interaction contrast (IC) in the data. That is,

$$\begin{aligned}\Delta^2_{\mathrm{L},\,\mathrm{R}} S_{\mathrm{RT}}(t) = {}& S(t)_{\mathrm{L}(\mathrm{l}),\,\mathrm{R}(\mathrm{l})} - S(t)_{\mathrm{L}(\mathrm{l}),\,\mathrm{R}(\mathrm{h})}\\ &- S(t)_{\mathrm{L}(\mathrm{h}),\,\mathrm{R}(\mathrm{l})} + S(t)_{\mathrm{L}(\mathrm{h}),\,\mathrm{R}(\mathrm{h})},\end{aligned}$$

where t is a specific value of the random variable **RT**. This defines the (population) interaction contrast, but other representations will often be useful in the proofs.

THEOREM 1. A. *Under the above conditions and assuming parallel processing, self-terminating (minimum time) processing will cause survivor function overadditivity*:

A-1. *If and only if*

$$\begin{aligned}\Delta^2_{\mathrm{L},\,\mathrm{R}} S_{\mathrm{RT}}(t) = \int_{t'=0}^{t} \int_{t_{\mathrm{R}}=t-t'}^{\infty} &\{[S_{\mathrm{L}(\mathrm{l})}(t-t' \mid \mathrm{T}_{\mathrm{R}} = t_{\mathrm{R}}, \mathrm{T}_{\mathrm{B}} = t')\\ &- S_{\mathrm{L}(\mathrm{h})}(t-t' \mid \mathrm{T}_{\mathrm{R}} = t_{\mathrm{R}}, \mathrm{T}_{\mathrm{B}} = t')]\\ &\times [f_{\mathrm{R}(\mathrm{l})}(t_{\mathrm{R}} \mid \mathrm{T}_{\mathrm{B}} = t') - f_{\mathrm{R}(\mathrm{h})}(t_{\mathrm{R}} \mid \mathrm{T}_{\mathrm{B}} = t')]\}\\ &\times f_{\mathrm{B}}(t')\, dt_{\mathrm{R}}\, dt' > 0.\end{aligned}$$

Now simplify the notation slightly (i.e., $\mathrm{T}' = t'$ *to simply* t'*, etc.).*

A-2. *If selective influence acts at the density one point crossing level, for any* $t' > 0$*, and*

$$\begin{aligned}\int_{t_{\mathrm{R}}=0}^{\infty} &[S_{\mathrm{L}(\mathrm{l})}(t-t' \mid t_{\mathrm{R}}, t') - S_{\mathrm{L}(\mathrm{h})}(t-t' \mid t_{\mathrm{R}}, t')]\\ &\times [f_{\mathrm{R}(\mathrm{l})}(t_{\mathrm{R}} \mid t') - f_{\mathrm{R}(\mathrm{h})}(t_{\mathrm{R}} \mid t')]\, dt_{\mathrm{R}} > 0,\end{aligned}$$

for all $t', t \geqslant 0$.

A-3. *If density one point crossing selective influence holds and* $S_{\mathrm{L}(\mathrm{l})}(t-t' \mid \mathrm{T}_{\mathrm{R}} = t_{\mathrm{R}}, t') - S_{\mathrm{L}(\mathrm{h})}(t-t' \mid \mathrm{T}_{\mathrm{R}} = t_{\mathrm{R}}, t')$ *is weakly monotonic increasing in* t_{R}*, for any* $t' > 0$.

A-4. *If the* R *and* L *parallel channels are conditionally (on* T_{b}*) stochastically independent of each other, though possibly dependent on* t_{B}*, with selective influence in force.*

B. *Any of the above conditions will cause mean reaction time overadditivity.*

Proof. A. First we derive the basic expressions. If processing is self-terminating, the reaction time random variable can be expressed as $\mathrm{RT} = \mathrm{MIN}(\mathrm{T}_{\mathrm{L}(i)}, \mathrm{T}_{\mathrm{R}(j)}) + \mathrm{T}_{\mathrm{B}}$ with the terms defined as above. Let $\Delta^2_{\mathrm{L},\,\mathrm{R}}$ represent the second-order partial difference over the two factor levels which defines the contrast operator (see, e.g., Townsend & Ashby, 1983). We now apply this operation to the observable reaction time survivor function

$$\begin{aligned}
&\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT}}(t)\\
&\quad = \Delta^2_{\mathrm{L,R}} P[\mathrm{MIN}(\mathbf{T}_{\mathrm{L}(i)}, \mathbf{T}_{\mathrm{R}(j)}) + \mathbf{T}_{\mathrm{B}} > t]\\
&\quad = \Delta^2_{\mathrm{L,R}} \left\{ S_{\mathrm{B}}(t) + \int_0^{t} P[\mathrm{MIN}(\mathbf{T}_{\mathrm{L}(i)}, \mathbf{T}_{\mathrm{R}(j)}) > t - t' \mid \mathbf{T}_{\mathrm{B}} = t'] \right.\\
&\qquad \left. \times f_{\mathrm{B}}(t')\, dt' \right\}\\
&\quad = \Delta^2_{\mathrm{L,R}} \int_0^{t} P[\mathrm{MIN}(\mathbf{T}_{\mathrm{L}(i)}, \mathbf{T}_{\mathrm{R}(j)}) > t - t' \mid \mathbf{T}_{\mathrm{B}} = t']\\
&\qquad \times f_{\mathrm{B}}(t')\, dt'\\
&\quad = \Delta^2_{\mathrm{L,R}} \int_{t'=0}^{t} \left[\int_{t_{\mathrm{R}} - t - t'}^{\infty} f_{\mathrm{R}(j),i}(t_{\mathrm{R}} \mid \mathbf{T}_{\mathrm{B}} = t') \right.\\
&\qquad \left. \times S_{\mathrm{L}(i),j}(t - t' \mid \mathbf{T}_{\mathrm{R}(j)} = t_{\mathrm{R}}, \mathbf{T}_{\mathrm{B}} = t')\, dt_{\mathrm{R}} \right] f_{\mathrm{B}}(t')\, dt'.
\end{aligned}$$

Simplifying the notation a bit and taking the second-order partial difference, this is

$$\begin{aligned}
&= \int_{t'=0}^{t} \left\{ \int_{t_{\mathrm{R}}=t-t'}^{\infty} [f_{\mathrm{R(l),l}}(t_{\mathrm{R}} \mid t')\, S_{\mathrm{L(l),l}}(t - t' \mid t_{\mathrm{R}}, t') \right.\\
&\quad - f_{\mathrm{R(l),h}}(t_{\mathrm{R}} \mid t')\, S_{\mathrm{L(h),l}}(t - t' \mid t_{\mathrm{R}}, t')\\
&\quad - f_{\mathrm{R(h),l}}(t_{\mathrm{R}} \mid t')\, S_{\mathrm{L(l),h}}(t - t' \mid t_{\mathrm{R}}, t')\\
&\quad \left. + f_{\mathrm{R(h),h}}(t_{\mathrm{R}} \mid t')\, S_{\mathrm{L(h),h}}(t - t' \mid t_{\mathrm{R}}, t')]\, dt_{\mathrm{R}} \right\} \cdot f_{\mathrm{B}}(t')\, dt'.
\end{aligned}$$

Under the assumption that the conditioned L survivor functions are invariant across the R factor levels this expression becomes

$$\begin{aligned}
&\int_{t'=0}^{t} \int_{t_{\mathrm{R}}=t-t'}^{\infty} \{ S_{\mathrm{L(l)}}(t - t' \mid t_{\mathrm{R}}, t')[f_{\mathrm{R(l),l}}(t_{\mathrm{R}} \mid t')\\
&\quad - f_{\mathrm{R(h),l}}(t_{\mathrm{R}} \mid t')] - S_{\mathrm{L(h)}}(t - t' \mid t_{\mathrm{R}}, t')[f_{\mathrm{R(l),h}}(t_{\mathrm{R}} \mid t')\\
&\quad - f_{\mathrm{R(h),h}}(t_{\mathrm{R}} \mid t')]\} \cdot f_{\mathrm{B}}(t')\, dt',
\end{aligned}$$

and under the additional assumption that the marginal R densities (i.e., marginal over t_{L} but conditional on $\mathbf{T}_{\mathrm{B}} = t'$) are invariant across the L factor levels and overadditivity is satisfied we obtain

$$\begin{aligned}
&\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT}}(t)\\
&\quad = \int_{t'=0}^{t} \int_{t_{\mathrm{R}}=t-t'}^{\infty} \{ [S_{\mathrm{L(l)}}(t - t' \mid t_{\mathrm{R}}, t') - S_{\mathrm{L(h)}}(t - t' \mid t_{\mathrm{R}}, t']\\
&\qquad \cdot [f_{\mathrm{R(l)}}(t_{\mathrm{R}} \mid t') - f_{\mathrm{R(h)}}(t_{\mathrm{R}} \mid t')]\}\, f_{\mathrm{B}}(t')\, dt_{\mathrm{R}}\, dt' > 0.
\end{aligned}$$

This is the result (A-1).

Basically (A-1) gives weak conditions for overadditivity to hold. Next, (A-2) introduces more bite when selective influence at the stronger level of density one point crossing holds. Part (A-2) claims that when t_{R} is integrated from 0 to ∞ and the result is positive, overadditivity holds in general. Write $S_{\mathrm{L(l)}}(t - t' \mid t_{\mathrm{R}}, t') - S_{\mathrm{L(h)}}(t - t' \mid t_{\mathrm{R}}, t') = G(t_{\mathrm{R}}, t, t')$ and $f_{\mathrm{R(l)}}(t_{\mathrm{R}} \mid t') - f_{\mathrm{R(h)}}(t_{\mathrm{R}} \mid t') = H(t_{\mathrm{R}} \mid t')$. Now $G > 0$ for all values of its argument according to hypothesis and $H(t_{\mathrm{R}} \mid t') \leqslant 0$ for $t_{\mathrm{R}} \leqslant t_*$, where t_* equals the density crossing point (for $0 < t < t_*$, $f_{\mathrm{R(l)}}(t \mid t') \mid < f_{\mathrm{R(h)}}(t \mid t')$ and for $t > t_*$, $f_{\mathrm{R(l)}}(t \mid t') > f_{\mathrm{R(h)}}(t \mid t')$; at $t = t_*$, $f_{\mathrm{R(l)}}(t_* \mid t') = f_{\mathrm{R(h)}}(t_* \mid t')$), and $H(t_{\mathrm{R}} \mid t') > 0$ for $t_{\mathrm{R}} > t_*$. Therefore the cumulative negativity is largest when t_{R} is integrated from 0 to $+\infty$. If this result is positive for any t, t', it will be positive for all other integration limits $t - t' > 0$. This concludes (A-2).

Clearly if $G(t_{\mathrm{R}}, t, t')$ is always (for all non-negative t, t') weakly monotonic increasing in t_{R}, then Proposition A-2 will ensue, thus proving (A-3).

Finally with conditioned stochastic independence of the parallel channels, we obtain from part (A-1)

$$\begin{aligned}
&\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT}}(t)\\
&\quad = \int_{t'=0}^{t} \int_{t_{\mathrm{R}}=t-t'}^{\infty} [S_{\mathrm{L(l)}}(t - t' \mid t') - S_{\mathrm{L(h)}}(t - t' \mid t')]\\
&\qquad \times [f_{\mathrm{R(l)}}(t_{\mathrm{R}} \mid t') - f_{\mathrm{R(h)}}(t_{\mathrm{R}} \mid t')] \cdot f_{\mathrm{B}}(t')\, dt_{\mathrm{R}}\, dt'\\
&\quad = \int_{t'=0}^{t} [S_{\mathrm{L(l)}}(t - t' \mid t') - S_{\mathrm{L(h)}}(t - t' \mid t')]\\
&\qquad \times [S_{\mathrm{R(l)}}(t - t' \mid t') - S_{\mathrm{R(h)}}(t - t' \mid t')]\, f_{\mathrm{B}}(t)')\, dt' > 0
\end{aligned}$$

by selective influence at the level of distribution ordering (density one point crossing is not required) and (A-4) is completed.

B. The overadditivity of the mean reaction time ensues immediately from integrating the survivor contrast from zero to infinity. This strategy follows because the integral of a survivor function *is* the mean (Townsend & Ashby, 1983, p. 170).

THEOREM 2. *Exhaustive* (*maximum processing time*) *parallel processes produce underadditivity in observed reaction times whenever self-termination* (*minimum processing time*) *produces overadditivity and vice versa.*

Proof.

$$\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT,EXH}}(t) = \Delta^2_{\mathrm{L,R}} P[\mathrm{MAX}(\mathbf{T}_{\mathrm{L}}, \mathbf{T}_{\mathrm{R}}) + \mathbf{T}_{\mathrm{B}} > t].$$

Now,

$$\begin{aligned} &P[MAX(\mathrm{T_L}, \mathrm{T_R}) + \mathrm{T_B} > t] \\ &\quad = P(\mathrm{T_L} > t - \mathrm{T_B} \text{ or } \mathrm{T_R} > t - \mathrm{T_B}) \\ &\quad = P(\mathrm{T_L} > t - \mathrm{T_B}) + P(\mathrm{T_R} > t - \mathrm{T_B}) \\ &\qquad - P(\mathrm{T_L} > t - \mathrm{T_B}, \mathrm{T_R} > t - \mathrm{T_B}). \end{aligned}$$

The first two terms vanish under application of $\Delta^2_{\mathrm{L,R}}$, assuming selective influence, while the third equals $P[\mathrm{MIN}(\mathrm{T_L}, \mathrm{T_R}) + \mathrm{T_B} > t]$. Hence, whenever the self-terminating (minimum time) case yields overadditivity, the exhaustive case yields underadditivity and conversely. ▌

Serial Process Models

The assumptions in the case of serial processing are analogous to those in the case of parallel processing. However, they are minimal for self-terminating serial processing because only a single item is processed on each trial. That is, serial processing predicates that an unprocessed item (or its level) cannot affect the processing time of another item processed earlier (Townsend, 1974).

In the self-terminating and the exhaustive processing cases, we assume that the base time, conditioned on a right or left processing time, is invariant across factor levels associated with that stimulus. With exhaustive serial processing, we assume, much as in the parallel processing theorems, that

$$\begin{aligned} S_{\mathrm{L}(i),\mathrm{l}}(t_\mathrm{L} \mid t_{\mathrm{R(l)}}, t_\mathrm{B}) &= S_{\mathrm{L}(i),\mathrm{h}}(t_\mathrm{L} \mid t_{\mathrm{R(h)}}, t_\mathrm{B}) \\ &= S_{\mathrm{L}(i)}(t_\mathrm{L} \mid t_\mathrm{R}, t_\mathrm{B}), \end{aligned}$$

that is, the processing survivor function (and therefore densities also) on, say, the left side, can depend on the base time as well as on the processing time on the other side but *not* directly on the factor level of the other side. It is further assumed that the marginal process densities are invariant across the other side's factor levels, that is, $\int_{t_\mathrm{L}=0}^{\infty} f_{\mathrm{R}(i)}(t_\mathrm{R} \mid t_\mathrm{L}) \, f_{\mathrm{L}(j)}(t_\mathrm{L}) \, dt_\mathrm{L}$ is, for $i = \mathrm{l}, \mathrm{h}$, invariant over $j = \mathrm{l}, \mathrm{h}$. It can be shown that these two conditions do not imply one another. Finally, it is assumed that the conditional survivor functions are ordered in their respective factors, e.g., $S_{\mathrm{L(l)}}(t \mid t_\mathrm{R}, t_\mathrm{B}) > S_{\mathrm{L(h)}}(t \mid t_\mathrm{R}, t_\mathrm{B})$ (the selective influence assumption operates at the survivor level) and similarly for the right side. The first theorem is on the behavior of survivor functions.

THEOREM 3. *Under the above prescribed conditions, serial models when self-terminating (minimum time) predict survivor function and mean reaction time factorial additivity.*

Proof. The density function on reaction time in the self-terminating case can be expressed as

$$\begin{aligned} f_{\mathrm{RT}}(t) = p \int_{t'=0}^{t} f_{\mathrm{L}(i)}(t') \, f_\mathrm{B}(t - t' \mid \mathrm{T_L} = t') \, dt' \\ + (1-p) \int_{t'=0}^{t} f_{\mathrm{R}(j)}(t') \, f_\mathrm{B}(t - t' \mid \mathrm{T_R} = t') \, dt' \end{aligned}$$

and of course $S_{\mathrm{RT}}(t) = \int_t^{\infty} f_{\mathrm{RT}}(t') \, dt'$, where p is the probability of processing the left stimulus first. It is immediately evident that subjecting the resulting expression to the second-order partial difference operator, $\Delta^2_{\mathrm{L,R}}$, renders a value of zero and therefore factorial additivity at the survivor function level. The mean contrast is shown to be zero by integration of the survivor contrast as in earlier proofs. ▌

Finally the results for serial exhaustive mean reaction times are considered.

THEOREM 4. *Exhaustive serial processing implies that survivor contrast for small t will be negative. It must become positive on intervals of larger t. More specifically, there exists a $t = t_*$ such that when $t < t_*$, t_* to be identified below, $\Delta^2_{\mathrm{L,R}} S_{\mathrm{R}t}(t) < 0$; that is, underadditivity holds in the observer survivor functions for sufficiently small t. Serial mean reaction contrast is always of zero contrast under these conditions.*

Proof. In the most general exhaustive case, the order of processing must be allowed to affect the reaction time (see, e.g., Townsend, 1974), in the present situation, at the distributional level. A subscript is affixed designating the stage of processing. Thus, $S_{\mathrm{R}_i(k)}$ is the survivor function for the right-hand signal for stage $i = 1, 2$ at factor level $k = \mathrm{l}, \mathrm{h}$. Let p equal the probability of processing the left side first. Then we may write

$$\begin{aligned} &\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT}}(t) \\ &\quad = \Delta^2_{\mathrm{L,R}} P(\mathrm{T_L} + \mathrm{T_R} + \mathrm{T_B} > t) \\ &\quad = \Delta^2_{\mathrm{L,R}} \left\{ p \int_{t_\mathrm{L}=0}^{t} \int_{t_\mathrm{B}=0}^{t-t_\mathrm{L}} S_{\mathrm{R}_{2(j)}}(t - t_\mathrm{L} - t_\mathrm{B} \mid t_\mathrm{L}, t_\mathrm{B}) \right. \\ &\qquad \times f_\mathrm{B}(t_\mathrm{B} \mid t_\mathrm{L}) \, f_{\mathrm{L}_{1(i)}}(t_\mathrm{L}) \, dt_\mathrm{B} \, dt_\mathrm{L} \\ &\qquad + (1-p) \int_{t_\mathrm{R}=0}^{t} \int_{t_\mathrm{B}=0}^{t-t_\mathrm{R}} S_{\mathrm{L}_{2(j)}}(t - t_\mathrm{R} - t_\mathrm{B} \mid t_\mathrm{R}, t_\mathrm{B}) \\ &\qquad \left. \cdot f_\mathrm{B}(t_\mathrm{B} \mid t_\mathrm{R}) \, f_{\mathrm{R}_{1(i)}}(t_\mathrm{R}) \, dt_\mathrm{B} \, dt_\mathrm{R} \right\}, \quad i, j = \mathrm{l}, \mathrm{h} \\ &\quad = p \int_{t_\mathrm{L}=0}^{t} \int_{t_\mathrm{B}=0}^{t-t_\mathrm{L}} [S_{\mathrm{R}_{2(l)}}(t - t_\mathrm{L} - t_\mathrm{B} \mid t_\mathrm{L}, t_\mathrm{B}) \end{aligned}$$

$$-S_{\mathrm{R}_{2(\mathrm{h})}}(t-t_{\mathrm{L}}-t_{\mathrm{B}} \mid t_{\mathrm{L}}, t_{\mathrm{B}})][f_{\mathrm{L}_{1(\mathrm{l})}}(t_{\mathrm{L}})-f_{\mathrm{L}_{1(\mathrm{h})}}(t_{\mathrm{L}})]$$
$$\cdot f_{\mathrm{B}}(t_{\mathrm{B}} \mid t_{\mathrm{L}})\, dt_{\mathrm{B}}\, dt_{\mathrm{L}}$$
$$+(1-p)\int_{t_{\mathrm{R}}=0}^{t}\int_{t_{\mathrm{B}}=0}^{t-t_{\mathrm{R}}}[S_{\mathrm{L}_{2(\mathrm{l})}}(t-t_{\mathrm{R}}-t_{\mathrm{B}} \mid t_{\mathrm{R}}, t_{\mathrm{B}})$$
$$-S_{\mathrm{L}_{2(\mathrm{h})}}(t-t_{\mathrm{R}}-t_{\mathrm{B}} \mid t_{\mathrm{R}}, t_{\mathrm{B}})]$$
$$\cdot[f_{\mathrm{R}_{2(\mathrm{l})}}(t_{\mathrm{R}})-f_{\mathrm{R}_{2(\mathrm{h})}}(t_{\mathrm{R}})]\, f_{\mathrm{B}}(t_{\mathrm{B}} \mid t_{\mathrm{R}})\, dt_{\mathrm{B}}\, dt_{\mathrm{R}}.$$

By the earlier hypotheses, the above two differences in survivor functions will be positive for all positive t. Also, it can be shown that a conditional ordering (for all values of the conditioning variables) implies a marginal ordering, which in turn implies that $f_{\mathrm{R(h)}}(t)$ density functions must be larger than $f_{\mathrm{R(l)}}(t)$ for small t, and the same goes for f_{L}. Suppose these values where the densities cross are $t_{1\mathrm{L}}, t_{1\mathrm{R}}$ for the two density difference, respectively, and take $t_* = \mathrm{MIN}(t_{1\mathrm{L}}, t_{1\mathrm{R}})$. Then it is implied that $\Delta^2_{\mathrm{L,R}} S_{\mathrm{RT}}(t) < 0$ for all $t < t_*$. The mean reaction time prediction is left to the reader. ▌

COROLLARY TO THEOREM 4. *In the presence of exhaustive processing, stochastic independence of all the processing and base times, along with selective influence, negative survivor function contrast holds for small times t. Mean contrast is zero.*

Proof. Left to reader (see text for a special case). ▌

The next theorem concerns survivor and mean contrast for an important type of coactivation model, the Poisson channel summation model (or superposition model of Schwarz, 1989; Colonius, 1990a; Diederich & Colonius, 1991).

THEOREM 5. *For the Poisson channel summation model*

A. *The survivor contrast is negative for small times t.*

B. *The survivor contrast is positive for large times t.*

C. *The mean contrast is positive.*

Proof. A. First, note that the Poisson channel summation model takes the sum of the individual rates as its rate. This implies that the survivor function will yield negative or positive contrast for a certain t, depending on whether it is a concave or convex function of its rate. Since the survivor contrast function is a well-behaved function of its rate parameter (in particular, with a continuous second derivative) the curvature will obey the sign of the second derivative. Hence we show that for small values of t, the second derivative is negative, implying negative contrast. Assigning the parameter u as rate, we see that

$$S_{\mathrm{L,R}}(t) = 1 - \sum_{j=k}^{\infty} \frac{t^j}{j!} u^j e^{-ut}$$

and

$$\frac{\partial^2 S_{\mathrm{L,R}}(t)}{\partial u^2} = -\sum_{j=k}^{\infty} \frac{t^j}{j!} u^{j-2} e^{-ut}[j(j-1) - 2jtu + t^2u^2].$$

It is obvious that the bracketed term approaches $j(j-1)$ as t approaches zero, proving our claim.

B. This case proceeds most readily using an alternate form of the survivor function. We write

$$S_{\mathrm{L,R}}(t) = \int_t^{\infty} \frac{(t')^{j-1}}{(j-1)!} u^j e^{-ut'}\, dt'.$$

The same logic as in (A) leads to

$$\frac{\partial S^2_{\mathrm{L,R}}(t)}{\partial u^2}$$
$$= \int_t^{\infty} \frac{(t')^j}{(j-1)!} e^{ut'} u^{j-2}[j(j-1) - 2jut' + u^2(t')^2]\, dt'.$$

The bracketed term in the integrand is positive for small t, goes negative for intermediate t, and then becomes positive for large t. This behavior, combined with the results of (A), implies that $S_{\mathrm{L,R}}(t)$ has early negative contrast, followed ultimately (for sufficiently large t) by positive contrast.

C. Positive mean contrast can be easily demonstrated from the form of the mean processing time, written in terms of left and right side processing rates,

$$E(T) = \frac{k}{u_{\mathrm{L,l}} + u_{\mathrm{R,l}}} - \frac{k}{u_{\mathrm{L,l}} + u_{\mathrm{R,h}}} - \frac{k}{u_{\mathrm{L,h}} + u_{\mathrm{R,l}}} + \frac{k}{u_{\mathrm{L,h}} + u_{\mathrm{R,h}}}.$$

where $u_{\mathrm{L,R}}$ is the left side rate set at low intensity, etc. ▌

APPENDIX B

Theorems on Capacity and Interactions and Predictions of Channel Summation Models

The first theorem relates Miller's inequality to capacity effects in parallel models, when comparing performance in one vs two (redundant) target conditions.

THEOREM 6. *Assume independent parallel processing with self-termination in the double stimulus condition.*

A. *If Miller's inequality is violated, that is, for some* $t = t_*$,

$$S_{\mathrm{L,R}}(t_*) < S_{\mathrm{L}}(t_*) + S_{\mathrm{R}}(t_*) - 1,$$

then $C(t_*) > 1$, *that is, the system is super capacity at any such* $t = t_*$.

B. *Suppose the parallel system is everywhere super capacity, in the sense that there exists a* $u > 1$ *such that* $C(t) \geqslant u$ *for all* $t > 0$. *Then there exists a* $t = t_*$ *such that Miller's inequality is violated for all* $t < t_*$.

Proof. A. When the violation of the inequality occurs, we have

$$S_{\mathrm{L,R}}(t_*) = \{S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*)\}^{C(t_*)} < S_{\mathrm{L}}(t_*) + S_{\mathrm{R}}(t_*) - 1.$$

We also know that for all t the following inequality holds:

$$S_{\mathrm{L}}(t) + S_{\mathrm{R}}(t) - 1 \leqslant S_{\mathrm{L}}(t) \times S_{\mathrm{R}}(t).$$

These two inequalities together produce

$$\{S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*)\}^{C(t_*)} < S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*).$$

Taking the logarithm of both sides of this inequality yields

$$C(t_*) \times \ln[S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*)] < \ln[S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*)]$$

and dividing both sides by $\ln[S_{\mathrm{L}}(t_*) \times S_{\mathrm{R}}(t_*)] < 0$, we get $C(t_*) > 1$, implying the existence of super capacity.

B. Suppose on the contrary that $C(t) \geqslant u > 1$ for all $t > 0$ and yet $[S_{\mathrm{L}}(t) \times S_{\mathrm{R}}(t)]^{C(t)} \geqslant S_{\mathrm{L}}(t) + S_{\mathrm{R}}(t) - 1$ for all $t > 0$; that is, Miller's inequality is satisfied. This is equivalent to the following expression:

$$C(t) \leqslant \ln[S_{\mathrm{L}}(t) + S_{\mathrm{R}}(t) - 1]/\ln[S_{\mathrm{L}}(t) \times S_{\mathrm{R}}(t)],$$

when t is small enough, that $S_{\mathrm{L}}(t) + S_{\mathrm{R}}(t) - 1 > 0$. As t approaches 0, the right hand expression smoothly approaches 1. This is a contradiction to the original assumption that $C(t) > u > 1$ for all t and shows that violations occur within an interval $(0, t_*)$. ∎

THEOREM 7. *Suppose processing is parallel and independent on the two signals and that the regular parallel model with self-termination and the channel summation model are based on the same underlying probability distributions for each signal. Then any such independent channel summation model is always less limited in capacity (more supercapacity) than the regular unlimited capacity parallel model for all counts k and times t.*

Proof. For any pair, $\mathrm{N_L} = n_{\mathrm{L}}$, $\mathrm{N_R} = n_{\mathrm{R}}$, $\mathrm{MAX}(n_{\mathrm{L}}, n_{\mathrm{R}}) \leqslant n_{\mathrm{L}} + n_{\mathrm{R}}$. Hence, for any positive (or zero) integer k, the probability that $\mathrm{MAX}(\mathrm{N_L}, \mathrm{N_R}) \leqslant k$ is always greater than (or equal to, but strictly greater for some n_{L}, n_{R} with nonzero probability) the probability that $\mathrm{N_L} + \mathrm{N_R} \leqslant k$. This implies that the survivor functions for the MIN processing time (indicating parallel minimum time and being identical to the MAX distribution function on count) are greater than the survivor function on the sum. Therefore, $S_{\mathrm{PAR}}(t) > S_{\mathrm{SUM}}(t)$ which can readily be shown to be equivalent to $C_{\mathrm{PAR}}(t) < C_{\mathrm{SUM}}(t)$, where $C_{\mathrm{PAR}}(t)$ represents the capacity function for the regular independent unlimited capacity, counting parallel model and $C_{\mathrm{SUM}}(t)$ is that for the channel summation model. ∎

Theorem 8 relates the most general channel summation models under inspection here, to Miller's inequality.

THEOREM 8. A. *Consider a channel summation model whose criterion is the same as that for the two simple signal conditions, whose criteria are also the same,* $n_1 = n_{\mathrm{R}} = k$. *Suppose this summation model is based on underlying, possibly dependent parallel processes observing the stipulation that there exists a* $t = t_*$ *such that for all* $t \leqslant t_*$,

$$\begin{aligned} &P[\mathrm{N_L} < k, \mathrm{N_R} < k, \mathrm{N_L} + \mathrm{N_R} \geqslant k, t] \\ &\quad = \sum_{n_{\mathrm{L}}=1}^{k-1} \sum_{n_{\mathrm{R}}=k-n_{\mathrm{L}}}^{k-1} P(\mathrm{N_L} = n_{\mathrm{L}}, \mathrm{N_R} = n_{\mathrm{R}}, t) \\ &\quad > P(\mathrm{N_1} \geqslant k, \mathrm{N_2} \geqslant k, t). \end{aligned}$$

Then such summation models violate Miller's inequality for a range of $0 < t < t_*$.

B. *A channel summation model based on parallel independent Poisson channels with criteria k in all three cases (left-alone, right-alone, and summed) predicts violation of Miller's inequality for all* $t < t_*$ *for sufficiently small* t_*.

Proof. A. The violated inequality can be expressed as

$$P(\mathrm{N_L} + \mathrm{N_R} \geqslant k, t) > P(\mathrm{N_L} \geqslant k, t) + P(\mathrm{N_R} \geqslant k, t) \quad \text{for some } t.$$

The left-hand side may be rewritten as

$$\begin{aligned} &P(\mathrm{N_L} + \mathrm{N_R} \geqslant k, t) \\ &\quad = P(\mathrm{N_L} \geqslant k, t) + P(\mathrm{N_R} \geqslant k, t) \\ &\qquad - P(\mathrm{N_L} \geqslant k, \mathrm{N_R} \geqslant k, t) \\ &\qquad + \sum_{n_{\mathrm{R}}=1}^{k-1} \sum_{n_{\mathrm{L}}=k-n_{\mathrm{R}}}^{k-1} P(\mathrm{N_R} = n_{\mathrm{R}}, \mathrm{N_L} = n_{\mathrm{L}}, t). \end{aligned}$$

Hence, the inequality becomes

$$\sum_{n_L=1}^{k-1} \sum_{n_R=k-n_L}^{k-1} P(\mathrm{N_R}=n_\mathrm{R}, \mathrm{N_L}=n_\mathrm{L}, t)$$
$$> P(\mathrm{N_L} \geqslant k, \mathrm{N_R} \geqslant k, t).$$

As $t \rightarrow 0$ the hypothesis implies that the stated violation will occur for a range of sufficiently small $0 < t < t_*$.

B. The proof for (B) easily comes about by showing that this model obeys the limiting property of (A) and is left to the reader. ▌

APPENDIX C

Development of Z-Tests for Survivor Function Interaction

We are interested in the double difference (i.e., the contrast) of survivor functions across the four factorial conditions. Time was divided into equal intervals of 10 msec with $n = 100$ being the total number of trials. At each interval the estimated survivor function can be considered as an estimate of the parameter p, probability, of a binomial random variable. Recall that $S(t) = 1 - F(t)$, where F is the cumulative distribution function. Then $\hat{S}(t) =$ (the number of observations greater than or equal to $t)/n$ is an estimate of the Bernoulli probability, p. If stochastic independence between trials is assumed, it is clear that the random variable will be distributed as $1/n \times \mathrm{K}(n, S(t))$, where $\mathrm{K}(n, S(t))$ is a binomial random variable with parameters n and $S(t) = p$. Since $n = 100$ in each estimated survivor function, the Normal approximation in the binomial random variable may be used, although we must expect some skewness for extreme probabilities $S(t)$. Hence, $\mathrm{S}(t)$ considered as a random variable is distributed approximately normally with mean $S(t)$, the true probability, and true standard deviation $S(t) \times F(t)/n$ (with the obvious estimate $\hat{S}(t) \times \hat{F}(t)/n$). The next step is representing the contrast of survivor functions as the double difference of the (approximately) normally distributed variables associated with the four factor conditions. This produces an overall Z-score, relative to a well-known estimate of the appropriate standard deviation, that can be utilized to test the contrast. The Z-score was thus defined as

$$Z = \Delta S^2_{\mathrm{L\,\&\,R}}(t)/\{\hat{S}_{\mathrm{ll}}(t) \times \hat{F}_{\mathrm{ll}}(t)/n_{\mathrm{ll}} + \hat{S}_{\mathrm{lh}}(t) \times \hat{F}_{\mathrm{lh}}(t)/$$
$$n_{\mathrm{lh}} + \hat{S}_{\mathrm{hl}}(t) \times \hat{F}_{\mathrm{hl}}(t)/n_{\mathrm{hl}} + \hat{S}_{\mathrm{hh}}(t) \times \hat{F}_{\mathrm{hh}}(t)/n_{\mathrm{hh}}\}^{1/2},$$

where n_{lh} is the total number of data points in the condition in which the left brightness is low and the right brightness is high and the other subscripts are similarly defined.

REFERENCES

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

Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, 21, 93–123.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, **93**, 154–179.

Bernstein, I. H. (1970). Can we see and hear at the same time? Some recent studies of intersensory facilitation of reaction time. *Acta Psychologica*, 33, 21–35.

Bjork, E. L., & Murray, T. J. (1977). On the nature of imput channels in visual processing. *Psychological Review*, **84**, 472–484.

Blake, R., Martens, W., Garrett, A., & Westendorf, D. (1980). Estimating probability summation for binocular reaction time data. *Perception & Psychophysics*, **27**, 375–378.

Blake, R., Martens, W., & Di Gianfilippo, A. (1980) Reaction times as a measure of binocular interaction in human vision. *Investigative Ophthalmology and Visual Science*, **19**, 930–941.

Blake, R., Martens, W., Garrett, A., & Westendorf, D. (1980). Estimating probability summation for binocular reaction time data. *Perception & Psychophysics*, **27**, 375–378.

Blake, R., Sloan, M., & Fox, R. (1981). Further development in binocular summation. *Perception & Psychophysics*, **30**, 226–276.

Brunk, H. D. (1960). *An introduction to mathematical statistics*. Boston: Ginn & Company.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523 545.

Burbeck, S. L., & Luce, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, **32**, 117–133.

Colonius, H. (1986). Measuring channel dependence in separate activation models. *Perception & Psychophysics*, 40, 251–255.

Colonius, H. (1988). Modeling the redundant signals effect by specifying the hazard function. *Perception & Psychophysics*, **43**, 604–606.

Colonius, H. (1990a). *Binocular reaction times to contrast increments: Probability and/or summation*? Manuscript submitted for publication.

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

Colonius, H., & Townsend, J. T. (1992). *State-space representation of general race and coactivation models*: *Prospects for testability*. Manuscript submitted for publication.

Colonius, H., & Vorberg, D. (1994). Distribution inequalities for parallel models with unlimited capacity. *Journal of Mathematical Psychology*, **38**, 35–58.

Colquhoun, W. (1975). Evaluation of auditory, visual and dual mode displays for prolonged sonar monitoring in repeated sessions. *Human Factors*, **17**, 425–437.

Diederich, A. (1992). *Intersensory facilitation*. Frankfurt am Main: Peter Lang.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, 39, 197–215.

Diederich, A., & Colonius, H. (1991). A further test of the superposition model for the redundant-signals effect in bimodal detection. *Perception & Psychophysics*, **50**, 83–85.

Donders, F. C. (1868). Over de snelheid van psychische processen. *Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechetsche Hoogeschool*, 1868–1869, *Tweede reeks*, II, 92–120. Translated by W. G. Koster, 1969, On the speed of mental processes. *Acta Psychologica*, **30**, 412–431.

Dzhafarov, E. N. (1992). The structure of simple reaction time to step-function signals. *Journal of Mathematical Psychology*, **36**, 235–268.

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

Dzhafarov, E. N., & Schweickert, R. (1994). Decompositions of response times: An almost general theory. Manuscript submitted for publication.

Egeth, H. E. (1966). Parallel vs. serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, 1, 245–252.

Egeth, H. E., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: Further evidence from subadditive effects of a visual quality. *Journal of Experimental Psychology: Human Perception & Performance*, 17, 550–559.

Egeth, H. E., & Mordkoff, T. J. (1991). Redundancy gain revisited: Evidence for parallel processing of separable dimensions. In G. R. Lockhead & J. R. Pomerantz (Eds.). *The Perception of Structure* (pp. 131–143). Washington, DC: American Psychological Association.

Egeth, H. E., Folk, C. L., & Mullin, P. A. (1988). Spatial parallelism in the processing of lines, letters, and lexicality. In B. E. Shepp and S. Ballesteros (Eds.), *Object perception: Structure and process*. Hillsdale, NJ: Erlbaum.

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception. Some results and methodological considerations. *Journal of Experimental Psychology Monographs*, 79 (Part 2).

Eriksen, C. W., Goettl, B., St. James, J. D., & Fournier, L. R. (1989). Processing redundant signals: Coactivation, divided attention, or what? *Perception & Psychophysics*, 45, 356–370.

Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception & Psychophysics*, 12, 278–286.

Estes, W. K., & Taylor, H. A. (1964). A detection method and probabilistic model for assessing information from brief visual displays. *Proceedings of the National Academy of Sciences*, 52, 446–454.

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.

Fournier, L. R., & Eriksen, C. W. (1990). Coactivation in the perception of redundant targets. *Journal of Experimental Psychology*, 16, 538–550.

Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

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

Goldstein, W. M., & Fisher, D. L. (1992). Stochastic networks as models of cognition: Deriving predictions for resource-constrained mental processing. *Journal of Mathematical Psychology*, 36, 129–145.

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

Grice, G. R., Canham, L., & Boroughs, J. M. (1984). Combination rule for redundant information in reaction time tasks with divided attention. *Perception & Psychophysics*, 35, 451–463.

Grice, G. R., Canham, L., & Gwynne, J. W. (1984). Absence of a redundant-signals effect in a reaction time task with divided attention. *Perception & Psychophysics*, 36, 565–570.

Gumbel, E. J. (1958). *Statistics of extremes*. New York: Columbia Univ. Press.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4, 11–26.

Hughes, R. C., Reuter-Lorenz, P. A., Nozawa, G., & Fendrich, R. (1994). Auditory-visual interactions in sensory-motor processing: Saccades versus manual responses. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 131–153.

Hughes, H. C., Reuter-Lorenz, P. A., & Nozawa, G. (1993). *Parallel and serial processes in the human oculomotor system: Bimodal integration and express saccades*. Unpublished manuscript

Kahneman, D. K. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall

Kantowitz, B., & Knight, J. L., Jr. (1974). Testing, tapping, time-sharing. *Journal of Experimental Psychology*, 103, 331–336.

Kendall, M. G., & Stuart, S. (1969). *The advanced theory of statistics* (Vol. 1), New York: Hafner.

Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics*, 15, 149–158.

Kinchla, R. A. (1977). Role of structural redundancy in the perception of visual targets. *Perception & Psychophysics*, 22, 19–30.

Kohfeld, D. L. (1969). Effects of the intensity of auditory and visual ready signals on simple reaction time. *Journal of Experimental Psychology*, 82, 88–95.

Krueger, L. E. (1978). A theory of perceptual matching. *Psychological Review*, 85, 278–304.

Krueger, L. E. (1984). Self-termination in same-different judgments: Multiletter comparison with simultaneous and sequential presentation. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 10, 271–284.

Legge, G. E. (1984). Binocular contrast summation. II. Quadratic summation. *Visual Research*, 24, 385–394.

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

McGill, W. J., & Gibbon, J. (1965). The general-gamma distribution and reaction times. *Journal of Mathematical Psychology*, 2, 1–18.

Meijers, L. M. M., & Eijkman, E. G. J. (1977). Distributions of simple RT with single and double stimuli. *Perception & Psychophysics*, 22, 41–48.

Meyer, D. E., & Schvaneveldt, R. D. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227–234.

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

Miller, J. (1986). Time course of coactivation in bimodal divided attention. *Perception & Psychophysics*, 40, 331–343.

Miller, J. O. (1988). Discrete and continuous models of information processing: Theoretical distinctions and empirical results. *Acta Psychologica*, 67, 191–257.

Miller, J. (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 160–169.

Mordkoff, J. T., & Egeth, H. E. (in press). Response time and accuracy revisited: Converging support for the interactive race model. *Journal of Experimental Psychology: Human Perception & Performance.*

Mordkoff, J. T., & Miller, J. (1993). Redundancy gains and coactivation with two different targets: The problem of target preferences and the effects of display frequency. *Perception & Psychophysics*, 53, 527–535.

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology*, 17, 520–538.

Neisser, U. (1967). *Cognitve Psychology*. Englewood Cliffs, NJ: Prentice Hall.

Nickerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement? *Psychological Reviews*, 80, 489–509.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitve Psychology*, 7, 44–64.

Nozawa, G. (1989). *A study of redundant target processing: Evidence for parallel processing*. Unpublished master's thesis, Purdue University, West Lafayette, IN.

Nozawa, G. (1992). *Stochastic Models of Human Information Processing*. Unpublished doctoral dissertation, Dartmouth College, Hanover, NH.

Nozawa, G, Reuter-Lorenz, P. A., & Hughes, H. C. (1993). *A model of bimodal integration and express saccades*. Unpublished manuscript.

Nozawa, G., Reuter-Lorenz, P. A., & Hughes, H. C (1994). Parallel and serial processes in the human oculomotor system: Bimodal integration and express saccades. *Biological Cybernetics*, 72, 19–34.

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). Potomac, MD: Erlbaum.

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

Piotrowski, S. M. (1983). *Orthogonal manipulations of stimulus contrast and discriminability: Implications for serial and parallel processing models.* Unpublished masters thesis, Purdue University, West Lafayette, IN.

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

Raab, D. H. (1961). *Statistical facilitation of simple reaction times.* Paper presented at the Division of Psychology, New York: New York Academy of Sciences.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53, 61–97.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66.

Schwarz, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, 46, 498–500.

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

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

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

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

Schweickert, R. (1992). Information, time, and the structure of mental events: A twenty-five year review. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 535–566). Cambridge, MA: MIT Press.

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

Schweickert, R., & Wang, Z. (1993). Effects on response time of factors selectively influencing processes in acyclic task networks with or gates. *British Journal of Mathematical and Statistical Psychology*, 46, 1–30.

Shapiro, R. G., & Krueger, L. E. (1983). Effect of similarity of surround on target letter processing. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 547–559.

Shaw, M. L. (1982). Attending to multiple sources of information. I. The integration of information in decision making. *Cognitve Psychology*, 14, 353–409.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis.* London: Chapman & Hall.

Smith, E. E. (1968). Choice reaction time: An analysis of the major theoretical positions. *Psychological Bulletin*, 69, 77–110.

Smith, P. L. (1990). A note on the distribution of response times for a random walk with Guassian increments. *Journal of Mathematical Psychology*, 34, 445–459.

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.

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*, 153 652–654.

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

Taylor, D. A. (1976). Holistic and analytic processes in the comparison of letters. *Perception & Psychophysics*, 20, 187–190.

Theios, J. (1973). Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 7, pp. 44–85). New York: Academic Press.

Townsend, J. T. (1968). Binocular information summation and the serial processing model. *Perception & Psychophysics*, 4, 125–128.

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* (pp. 133–168). New York: Wiley.

Townsend, J. T. (1981). *In pursuit of nonadditive reaction time theories.* Paper presented in Symposium on Modeling of Complex Cognitive Processes, at the Fourteenth Annual Mathematical Psychology Meetings, Santa Barbara, CA.

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

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

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

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan and F. Restle (Eds.), *Cognitive theory* (Vol. III, pp. 200–239). Hillsdale, NJ: Erlbaum.

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

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

Townsend, J. T., & Piotrowski, S. M. (1981, November). *Parallel systems and factorial additivity of mean reaction time.* Paper presented at the Psychonomic Society Meeting, Philadelphia, PA.

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

Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, 33, 309–327.

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

Townsend, J. T., & Van Zandt, T. (1990). New theoretical results on testing self-terminating vs. exhaustive processing. In H. G. Geissler and H. Schroeder (Eds.), *Proceedings of the International Fechner Symposium.* Amsterdam: North Holland.

Ulrich, R., & Giray, M. (1986). Separate-activation models with variable base times: Testability and checking of cross-channel dependency. *Perception & Psychophysics*, 34, 248–254.

van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel processing of redundant targets in simple visual search tasks. *Psychological Research*, 45, 235–254.

Van Zandt, T., & Townsend, J. T. (1993). Self-terminating vs exhaustive processes in rapid visual, and memory search: An evaluative review. *Perception & Psychophysics*, 53, 563–580.

Welford, A. T. (Ed.). (1980). *Reaction times*. New York: Academic Press.

Westendorf, D., & Blake, R. (1988). Binocular reaction times to contrast increments. *Vision Research*, 28, 355–359.

Received: March 6, 1995

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