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

Vol. 33, No. 3, September 1989
Printed in Belgium

Toward the Trichotomy Method of Reaction Times: Laying the Foundation of Stochastic Mental Networks

JAMES T. TOWNSEND AND RICHARD SCHWEICKERT

Department of Psychological Sciences, Purdue University

This article and the following one by Schweickert and Townsend develop a theory of mental architecture and function which leads to an experimental method of identifying that architecture. Our approach follows the tradition established by F. C. Donders (1868, *Archiv für Anatomie und Physiologie*, 657–681), W. Wundt (1896, *Grundriss der Psychologie*, Leipzig: Engelman) and others in the nineteenth century of postulating a class of architectures that describe how separate mental processes may be linked together. Then further assumptions are made about how experimental manipulations can affect those processes. Finally, predictions are derived that permit conclusions to be drawn about the form of the specific architecture engaged in a certain task. The first paper establishes terminology, postulates, and the general approach, and compares these with those of other theories and methods. Certain theoretical results are developed that set the stage for application. The second paper implements the theory by proving theorems that state how the experimental results obtained with the proposed methodology can be used to identify the underlying architecture. Further motivation and real and imaginary examples will also be presented there.

THEORETICAL FOUNDATIONS: A BRIEF AND INFORMAL HISTORY

Rather than jumping into the detailed terminology too quickly, it may be helpful to provide some motivation by outlining the primary historical antecedents. This will be done informally with more exact specification accomplished after the theoretical framework is set up. Donders (1868) is usually given credit for inventing the "method of subtraction" (e.g., see Boring, 1957). It is assumed that mental operations take place in a nonoverlapping series of processing durations so that the overall reaction time random variable is just the sum of the intervening random intervals. The average reaction time is then the expected value of the sum. Donders assumed that by simplifying the mental task the experimenter could 'subtract' a mental process or conversely could make the mental task more complex by 'adding' a mental process to the overall set of mental operations. Finally, by subtracting the average reaction time for the simpler task from that of the more complex task, the

Helpful comments by the referees, Hans Colonius and Ewart Thomas, and the Action Editor, Dirk Vorberg, led to the improvement of this manuscript. This work was supported by NSF Grant 8710163 to Townsend and NIMH Grant MH38675-01 to Schweickert. Thanks are due to Kilsoon Cummings for her help in typing this and the following paper.

0022-2496/89 $3.00

experimenter could estimate the mental duration required to perform the deleted operation. Although the assumptions of this method are indeed very strong, the method continues to be useful in certain circumstances (e.g., McGill, 1963; Sternberg, 1966; Theios, 1973; Ashby, 1982; Ashby & Townsend, 1980; Vorberg, 1981; Kohfeld, 1969; Kohfeld, Santee, & Wallace, 1981a,b; Gottsdanker & Shragg, 1985; Meyer & Schvaneveldt, 1971; Posner, 1978).

Reaction time continued to be employed over the next 100 years, and is still a vital tool in the investigation of elementary cognitive processes (Luce, 1986). The next major innovation in the tradition of Donders was the "additive factor method", introduced by Sternberg (1969). In this method, seriality is asserted just as in the Donders approach but experimental factors are sought which may affect separate processes of psychological interest. This notion is known as "selective influence". Some technique is used to assess whether the mean reaction times across the various factorial conditions interact or not. Analysis of variance is a common way of doing this. If the factors do not interact, they are said to be "additive" and it is concluded that the hypothetical psychological processes do indeed exist and are serially arranged. If the mean reaction times do interact, then it is concluded that both factors affect the same process or "stage". Hence, either the two hypothesized stages do not exist as separate psychological mechanisms, or the wrong experimental factors have been applied. Although some comments are given by Sternberg (1969) as to other architectural possibilities the clear target is the class of serial systems. Thus, seriality is in principle untestable within this method. Of course, the experimenter could instead, when faced with a factorial interaction, conclude that seriality failed, having postulated that selective influence holds. The method never seems to have been applied in that fashion. Observe also that a move has been made from Donders' attempt to assess the actual processing time consumed by certain psychological processes, or at least their averages, to a more qualitative question concerning the architecture. That is, based on an a priori postulate of seriality, Sternberg poses the question as to the existence of separate processes that might be associated with each of the experimental factors. General discussions of the additive factor method may be found in Theios (1973), Pachella (1974), Taylor (1976), Pieters (1983), Townsend (1984).

Townsend (1974a) took a somewhat different tack. Accepting the association of seriality with mean reaction time additivity, he undertook to apply factorial thinking to parallel processes. This was accomplished by postulating selective influence in parallel structures and mathematically investigating their behavior. The result was a prediction of factorial "subadditivity" in the class of independent parallel models. That is, after one factor was used to prolong reaction times, it was predicted that the other factor would have a diminished effect in terms of further prolongation. Recall that in Sternberg's system the two hypotheses were "serial with the hypothesized stages or processes" vs "the nonexistence of separate psychological stages affected by the two experimental factors". The new strategy was "additive implies seriality" vs "subadditivity implies parallelity" vs "other architectures or failure of selective influence". Thus, the qualitative sphere of identified architectures began to be enlarged. The class of parallel models was further extended by Townsend and Ashby (1983; Chap. 12). The latter work also firmed up the conditions under which additivity within serial systems could or could not occur.

Schweickert (1978, 1980, 1983a) developed a general theory along with a methodology for dealing with the entire class of so-called PERT networks, again based on a postulate of selective influence. These will be specified further below but suffice it to say now that virtually all architectures which assumed feedforward processing (i.e., no feedback loops) together with nonoverlapping durations of sequentially arranged processes were covered by this framework. It was termed "latent network theory" (Schweickert, 1983b). Its main limitation was that most results were based on an assumption of determinism: Each process always consumed exactly the same amount of time from trial to trial, although different processes could take different times to be completed. Considering stochastic rather than deterministic networks, Schweickert (1977, 1982) derived estimators and bounds on their bias of an important quantity in his theory, the coupled slack. Fisher and Goldstein (1983), Fisher, Saisi, and Goldstein (1985), and Goldstein and Fisher (in press) have developed algorithms for predicting the moments and completion time distributions for such architectures. The latter methods are especially useful when the individual processing distributions are within the general gamma family. Vorberg (1988) has recently suggested a new recurrence relation that facilitates calculations within that family. However, at this point there were still no general theorems available regarding the type of behavior to be expected with factorial methods in the context of the more complex latent networks.

Townsend and Schweickert (1985) generalized Townsend's (1974), Schweickert's (1978), and Townsend and Ashby's (1983) results on parallel and concurrent processes, still finding predictions of subadditivity. They also discovered that when two serial manipulated processes were in parallel with a third process, the prediction was of superadditivity. That is, it was predicted that after prolongation of reaction time by one factor, the other would have an enhanced effect on reaction time. Schweickert, Townsend, and Fisher (1987) further generalized the class of random latent networks which can be classified by factorial methods and several of their results are contained in the present paper. Observe that all this work assumes that selective influence is in force.

Some Motivation for the Present Class of Mental Networks

The train of the above research is clearly in the direction of generalizing the classes of architectures that can be identified by factorial experiments, assuming selective influence. Of course, the class of networks now covered does not exhaust all possible kinds of mental architectures or their functioning. As noted above, feedback systems are not presently included and networks which permit the temporal overlap of sequentially arranged processes are also not covered. The latter have been termed "continuous flow" systems (e.g., Eriksen & Hoffman, 1972). Taylor

(1976), McClelland (1979), Meyer, Yantis, Osman, and Smith (1985), and Miller (1982a, b, 1988) have contributed to the understanding of sequentially arranged, continuous flow processing systems, but little in the way of general theoretical results is yet available. For instance, it seems a certainty that, within specified tolerances, continuous flow networks will mimic the present discrete flow networks but this remains to be proven rigorously.

Systems based entirely on holographic-like behavior, without any communication among separate processes, are omitted from our purview. For instance, a single processor engaged in Fourier analysis or wavelength interference production does not fall within our theory. However, our results may relate to the global behavior of a system that posits one or several such processes interacting with other processing units.

Basically our stance is that the present theory and related methodology enclose a broad class of potential mental architectures that seem likely to be of interest to cognitive and experimental psychologists. Further, it will become clear that the much studied cases of serial and parallel processing are special instances of the present scheme. And these are still ubiquitous in the cognitive and information processing literature as a scan of the contemporary journals in those disciplines will document. Also, the networks developed and analyzed in this paper may approximate other classes of networks as well, as intimated above in the case of continuous flow systems. The present "theory" is, in a sense, a metatheory which permits the classification of particular networks within its system of taxonomy. It will be shown in the sequel, moreover, that there exist ways of testing the entire theory. That is, empirical results are possible that could falsify all networks within the present theory (at least for that particular data set).

We also believe that theory and experimentation go hand in hand so that as the present methodology becomes disseminated, the experimental milieu will grow in such a fashion as to exploit it. However, there are already a few examples of application, which will be pointed out in the sequel to this paper.

Basic Concepts of Stochastic Discrete Mental Networks

We have been using terms somewhat loosely in order to keep the introduction simple. We continue to employ words like "architecture" and "system" informally to denote organizations of cognitive mechanisms and operations. However, we need to become more formal with terminology that plays a strategic role in the theory development. We will require definitions of the types of architectures covered here and various notions that relate the theory to experimentation. The major concepts will be set aside from the main text but minor items will be defined less formally. A more complete presentation is available in Schweickert (1978, 1982).

Definition 1. A. A *directed graph* is a nonempty set V of vertices, a set A of arcs disjoint from V, and a mapping m of A into $V \times V$. The mapping m associates with every arc x in A two vertices $\langle x', x'' \rangle$, the starting and finishing vertices, respectively, of x.

B. A *path* from vertex a to vertex b is an alternating sequence of vertices and arcs, a, x_1, v_1, x_2, v_2, ..., x_m, b, where a is the starting vertex of arc x_1, b is the finishing vertex of x_m, and v_i is the ending vertex of arc x_i and also the starting vertex of arc x_{i+1} for $1 \leqslant i < m$. A path may consist of exactly one vertex.

C. A *cycle* is a path from a vertex a to itself, such that the path has at least one arc.

D. A *directed acyclic network* is a directed graph with a variable associated with each arc and with no cycles.

In pictorial illustrations, the binary relations are represented by directed arcs between the vertices. An intuitive way to think of a directed network is as a graph connecting the vertices, with arrows (directed arcs) denoting the direction of information flow, or the order of processing. This leads to the next definition.

DEFINITION 2. A *mental process* is a psychological entity identified with an arc in the graph of a network.

We shall typically use the simpler term "process". This term seems preferable in general contexts to the common term "stages" (cf. Sternberg, 1969) because the latter typically connotes seriality and it has also been used to denote an epoch between completions of items (Townsend & Ashby, 1983). Miller (1988) distinguishes "stages" and "processes" with the former denoting the locus of the particular mechanism and the latter referring to the actual operations. The present investigation uses "process" for both, although the locus in the overall architecture carries the major theoretical force here. The next definition establishes the overall class of architectures to be studied. Recall that a partial order on a set of objects A is a binary relation R that is reflexive (if x is in A then x has the relation to itself, that is xRx), transitive (if x, y, and z are in A and if xRy and yRz then also xRz), and antisymmetric (if x and y are in A and both xRy and yRx then $x = y$).

DEFINITION 3. A *stochastic, discrete mental process network* is a partial order on a set of processes with the properties that

(i) All processes preceding any specific process must be completed before the specified one begins.

(ii) There exists a joint density on the processing times of all processes in the network.

In a partially ordered set of process can be represented as an arc in a directed, acyclic network. The start of a process x is known as a starting vertex x' and its terminal vertex is denoted x''. A process x is adjacent to a process y if the terminal vertex of x is the same as the starting vertex of y or vice versa. The duration of a process is a nonnegative random variable $D(x)$, assumed to have finite mean, and a value taken on by $D(x)$ is denoted $d(x)$. The *duration of a path* is the sum of the durations of all the processes on the path. Note that a path may consist of exactly one vertex, in which case its duration is 0. Since a path duration is the sum of a number of random variables, the path duration itself is also a random variable.

It is interesting that (i) sets up the discrete flow concept but it also has stronger implications of exhaustiveness on the part of preceding processes. One could equally well define a self-terminating rule which specifies a certain subset of preceding processes that must be done before the specified one starts. That too would be a type of discrete flow network. It seems that a theory parallel to the present one could be constructed based on such notions and, indeed, this may be desirable for certain kinds of architectures; this lies beyond the present scope.

The joint distribution, of course, establishes the random quality of the mental networks. For purposes of relating the present work to other results in applied mathematics and engineering, it should be noted that the above definition specifies "stochastic PERT networks." We shall refer simply to "networks" rather than the entire term of the above definition for simplicity.

The next concepts are so important that we take special care to delineate them.

DEFINITION 4. A. If there is a path from the finishing vertex of x to the starting vertex of y, or vice versa, then x and y are *sequential.* Otherwise they are *concurrent.*

B. If x and y are sequential and another process is on a path with y if and only if it is on a path with x, then x and y are in *series.*

C. If x and y are concurrent and they have the same starting vertex and the same terminal vertex, then they are said to be in *parallel.*

It may seem paradoxical that the definition for "serial" permits x and y to have other processes before x, between x and y, and after y, yet the definition of "parallel" does not permit x to be preceded or followed by z with both concurrent with y. This apparent incongruity is the outcome of the graph theory definitional structure, which entails that the present serial and parallel structures are duals of one another.

A "source" is a vertex preceded by no arc, and a "sink" is a vertex followed by no arc. In a typical psychological task, a stimulus is presented at a source, and a response is made at a sink. We will assume here that there is one source, o, and one sink, r. Networks with multiple sources and sinks are discussed in Schweickert (1978, 1980, 1983b).

Important concepts for the analysis of networks are "critical path" and "slack" whose definitions follow. Intuitively, the critical path is the longest path through the network and the slack refers to how much a process needs to be prolonged before it affects the start time of a process y, or the overall finishing time.

DEFINITION 5. A. A *critical path* of a network on a trial is a path from the source to the sink whose duration is maximal over all other paths through the network. The duration of the critical path is the response time for that trial.

B. The *slack* $s(xy')$ from process x to process y is the longest amount of time by which x can be prolonged without delaying the onset of y.

C. The *total slack* $s(xr)$ for a process x is the longest amount of time by which x can be prolonged without delaying the response made at r.

Clearly the response time on any given trial is the sum of times of the arcs in the critical path on that trial.

Computationally, the slack from x to y, $s(xy')$, equals the difference between the duration of the longest path from the source o to the starting vertex y' of y and the duration of the longest path containing x from o to y'. That is, letting $d(a, b)$ stand for the duration of the longest path from vertex a to vertex b, we find that $s(xy') = d(o, y') - d(o, x') - d(x) - d(x'', y')$ (Schweickert, 1978). Note that if d has one argument, the duration of a process is denoted, while if d has two arguments, the duration of a path is denoted.

Similarly, the total slack, $s(xr)$, equals the difference between the duration of the longest path from o to r, and the duration of the longest path from o to r containing x. Thus, $s(xr) = d(o, r) - d(o, x') - d(x) - d(x'', r)$. It turns out that a quantity $s(yx'')$ is valuable in theoretical work: $s(yx'') = s(yr) - s(xr) + s(xy')$, even though x precedes y. It may be thought of as the slack from y to x in the network derived from the original task network by reversing the directions of all the arrows. In the new network, y would precede x (for more detail see Schweickert, 1978).[1]

When the durations of the processes are random variables, then so are the path durations and the slacks. If we let the random duration of the longest path from vertex a to vertex b be denoted $D(a, b)$, then the following random variables are the slacks of interest. For processes x and y,

$$S(xr) = D(o, r) - D(o, x') - D(x) - D(x'', r)$$

and

$$S(yr) = D(o, r) - D(o, y') - D(y) - D(y'', r).$$

If process y follows process x,

$$S(xy') = D(o, y') - D(o, x') - D(x) - D(x'', y').$$

[1] We define the *coupled slack from* x *to* y *to be*

$$k(xy) = s(xr) - s(xy').$$

Coupled slack is important in the deterministic situation, where the process durations are real numbers rather than random variables. If the processes are prolonged by amounts which are not too small, then the combined effect on reaction time of prolonging two sequential processes x and y is the sum of the effects of prolonging them individually plus the coupled slack, $k(xy)$. The details are beyond the scope of this paper; see Schweickert (1978).

We have developed the concepts for the architectures. Now we need structure to implement the factorial effects on the response times.

Selective Influence and Pure Insertion

Although this and the following paper are largely concerned with influences of experimental variables on a fixed network, it is important to place our concepts within a reasonably general framework, Previous treatments (our own included) prove somewhat limited when the province of general mental architectures is entered. The framework with regard to factorial influences developed here is more rigorous and general than heretofore, but we do confine it to stochastic discrete mental networks as defined above.

One topic that requires discussion is the relation of the factorial methods to the method of subtraction, especially as regards the basic definitions and assumptions. The method of subtraction postulates the axiom of pure insertion, which ensures that a process can be added to or deleted from the overall architecture (e.g., Ashby & Townsend, 1980). The axiom of selective influence, on the other hand, assumes that a specific experimental variable influences one and only one process. Definitions given in terms of the concepts of the architecture are required.

DEFINITION 6. *Pure insertion* is the addition of a process to a network, by either joining two vertices with an arc or replacing a single arc with two arcs in series. Thus, the graph of a network is enlarged by a single arc and the specification of the joint density of processing times by a random vector is incremented by a single random variable.

Ordinarily the additional arc will have its terminal vertex attached to a vertex already in the network. Often both its vertices will be attached to vertices of arcs already in the network. Note that although there is an additional random variable represented in the network specification, it may or may not have an effect on the response times. Obviously, the interest of the psychologist is in cases where it does. The next definition is on selective influence.

DEFINITION 7. *Selective influence* occurs when an experimental variable (factor) has a direct effect on the processing time distribution of exactly one process.

Thus, some aspect of the distribution on the processing time random variable associated with the specific process (arc in the network) is altered. This word "direct" in the preceding sentence is there for good reason. It is possible in a network which permits stochastic dependencies among its processes that an indirect influence on process A from an experimental variable which directly affects another process B could occur. This is referred to as an "indirect nonselective influence" (Townsend, 1984). It will be returned to below. Note that we here confine the notion of selective influence to a single process. Future work may find it useful to extend this concept to influence of several processes.

We further assume that experimental factors can be chosen so that the process random variable is lengthened or shortened in some well defined sense. Townsend and Ashby (1978; 1983) found that some distributional influences are stronger than others. Townsend (1986) provided more results along that line and constructed an implication graph which shows the dominance relations among these. One of the weakest influences is that the factor orders the means of the processing times of the individual process. An extremely strong influence is in force if the likelihood ratio of the density function of the process, under two different values of the factor, is monotonic (Townsend & Ashby, 1983; Townsend, 1986).

Clearly we wish to have a factorial influence that is sufficiently strong to generate predictions in complex architectures, but not so strong as to be highly unrealistic. An intermediate degree of influence is when the factor orders the distribution functions of the individual processing time random variable (that is, of a single arc). This implies the ordering of the means but not vice versa. It turns out to be strong enough to permit the empirical classification of networks through mean response times that we seek. Schweickert's latent mental network theory (1978, 1982, 1983b) assumed that changing a factor simply added a duration to the process to be influenced. In the present case, where durations are taken to be stochastic, the increment would itself have to be a random variable. It turns out that these two modes of influence are basically the same thing: Adding a random variable to a distribution implies that the distributions are ordered. Furthermore, if two processing time distributions are ordered (by action of the factor in the present application) then there exists a positive random variable which can be added to the random variable of the faster distribution to produce the slower.

More precisely, let the distribution of the duration of process x when the factor is at level i be denoted $F_i(t)$.

DEFINITION 8. A factor with n levels is said to *selectively influence* process x by *ordering its distribution* if for all t, $F_i(t) \geqslant F_j(t)$ whenever $i<j$, for $1 \leqslant i<j \leqslant n$.

Now let $D_i'(x)$ be the duration random variable of process x when the factor affecting x is at level i. Observe that $D_i'(x)$ has the same marginal distribution as the original duration random variable for the process x, $D_i(x)$. However, $D_i'(x)$ and U_{ij} may not be independent. From here on, we shall often abuse the notation for simplicity of expression by writing $D_i'(x) = D_i(x)$, that is, dropping the "prime." Observe that the following interpretation is manifestly *not* implied: $P[D_j(x) > D_i(x)] = P[D_i(x) + U_{ij} > D_i(x)] = 1$.

DEFINITION 9. The experimental factor is said to *selectively influence process x by incrementing its duration* if for all levels i and j of the factor with $1 \leqslant i<j \leqslant n$, $D_j(x) = D_i'(x) + U_{ij}$, where U_{ij} is a nonnegative random variable and $D_i'(x)$ is a random variable with the same marginal distribution as $D_i(x)$.[2]

[2] The use of $D_i(x)$ rather than $D_i'(x)$ seems neither more nor less innocent than the use of a single symbol μ to represent the common regression coefficient in the usual two-factor design with mean values $E(X_{ij}) = \mu + \mu_i + \mu_j$ (e.g., Wilks, 1962). This expression suggests, of course, an additive decomposition of underlying random variables, $X_{ij} = X + X_i + X_j$.

The following theorem shows that these two types of selective influence are effectively the same. In this and the subsequent work, it is not necessary to assume that the distribution functions are strictly increasing or even continuous, although it is required that means exist. However, we have decided to present the results based on the presence of densities for the benefit of a wider audience of behavioral scientists.[3]

THEOREM 1. *Let the distribution function of the duration of process x when a factor is at level i be $F_i(t)$, for $1 \leqslant i \leqslant n$, and suppose F_i^{-1} exists for $1 < i \leqslant n$. A factor orders the distribution function of process x if and only if it increments the duration of process x.*

Proof. A. Without loss of generality, suppose the levels of the factor are 1 and 2. Suppose $F_1(t) \geqslant F_2(t)$ for all $t \geqslant 0$. If T is a random variable with distribution function $F(t)$, then $F(T)$ is a random variable with the uniform distribution; that is, $P(F(X) \leqslant x) = x$, where $0 \leqslant x \leqslant 1$. Define

$$U = F_2^{-1}[F_1(T_1)] - T_1.$$

Then U is positive because $F_1(t) \geqslant F_2(t)$ and $P(T_1 + U \leqslant t) = P[F_2^{-1}(F_1(T_1)) \leqslant t] = P[F_1(T_1) \leqslant F_2(t)] = F_2(t)$.

[3] For instance, Hans Colonius has suggested the following proof for part A of Theorem 1, which does not assume strictly monotonic F's:

Proof of "ordering implies incrementing." Assume $F_2(t) \leqslant F_1(t)$ for all t; define for $0 < y < 1$

$$F_i^{-1}(y) = \inf\{t: F_i(t) \geqslant y\} \quad (i = 1, 2);$$

obviously, F_i^{-1} are nondecreasing functions and

$$F_i(F_i^{-1}(y)) \geqslant y.$$

Then $y \leqslant F_i(t_0)$ implies $F_i^{-1}(y) \leqslant F_i^{-1}(F_i(t_0)) \leqslant t_0$ and $F_i^{-1}(y) \leqslant t_0$ implies $y \leqslant F_i(F_i^{-1}(y)) \leqslant F_i(t_0)$ for any real t_0, i.e., $y \leqslant F_i(t)$ is equivalent to $F_i^{-1}(y) \leqslant t$. Let Y be uniformly distributed on $(0, 1)$. Then $F_i^{-1}(Y)$ are nonnegative random variables and

$$\Pr(F_i^{-1}(Y) \leqslant t) = \Pr(Y \leqslant F_i(t)) = F_i(t) \qquad \text{for all } t.$$

Moreover, $F_2(t) \leqslant F_1(t)$ for all t implies

$$F_1^{-1}(y) \leqslant F_2^{-1}(y) \qquad \text{for all } y.$$

Now set

$$D_1' = F_1^{-1}(Y), \qquad D_2 = F_2^{-1}(Y), \qquad \text{and} \qquad U_{12} = F_2^{-1}(Y) - F_1^{-1}(Y);$$

it follows that U_{12} is nonnegative, D_1' has distribution function F_1, and

$$\Pr(D_1' + U_{12} \leqslant t) = \Pr(D_2 \leqslant t) = F_2(t).$$

B. Now suppose $T_2 = T_1 + U$. Then,

$$P(T_1 + U \leqslant t) = P(T_1 \leqslant t - U) = \int_0^t P(T_1 \leqslant t - u \mid U = u) f_u(u)\, du$$

$$\leqslant \int_0^t P(T_1 \leqslant t \mid U = u) f_u(u)\, du = P(T_1 \leqslant t). \qquad \text{Q.E.D.}$$

The above treatment provides a new view of selective influence and pure insertion. Our definition of selective influence combines its locus (each experimental factor influences a single process; different factors influence different processes) and its effect (distribution ordering of affected processes) in the presence of experimental factors. Pure insertion specifies an effect with or without designation of a possible locus. Ordinarily, the experimenter will have some idea as to the potential networks and positions into which a process will be purely inserted. However, that may not always be the case. The experimental effect of a pure insertion is to add a random variable somewhere in the network. Note that both pure insertion and selective influence could be generalized to include several processes rather than a single process. Such a theory and its related methodology would undoubtedly be more complex than the present one, although novel concepts should not arise in the simple cases of parallel and serial architectures.

Next, it should be observed that if selective influence is in force at a certain locus at the level of distribution ordering, then it is equivalent to the pure insertion of a process just before or just after the selectively influenced process (i.e., at the locus affected by the experimental factor). There are several things that should be said about this equivalence.

First, while many purely inserted processes might be expected to be stochastically independent of the surrounding processes, it seems unlikely that the incrementing random variable which comes with selective influence at the distribution ordering level will often be independent of the random variable representing the initial process duration. However, if an inserted process uses important information from the preceding process, or feeds such information into the subsequent process, then a stochastic dependence may result. It seems likely that most experimental situations will make it clear whether a pure insertion or a selective influence is more probable.

Nevertheless, there do exist some cases where specification of which type of effect is occurring may be impossible in principle. It appears that these are mostly of philosophical interest in that it does not seem to matter much which one is present. Thus, consider the situation when a process is selectively influenced by increasing the number of items it has to process, or some analogous operation. Each new item leads to a new random variable representing the processing time of the item by the original process. Obviously, there is a perfectly equivalent system—that is, equivalent both mathematically and with regard to the experimental milieu—which purely inserts a new process for each additional item. Each process works only on a single item for which it is responsible. Note that not only can the mathematical description of the two systems be equivalent, but it is questionable if any factorial manipulation can discriminate them. Any factor that influences the process in the first system would presumably influence all the processes in the second system in the same way. The usual short term memory search and brief visual display search experiments are of this type, of course (Sternberg, 1966; Atkinson, Holmgren, & Juola, 1969; Townsend & Roos, 1973). In any case, in most systems there appears to be little danger from a lack of discriminability and in the latter example, it is not clear that any harm comes from the indicriminability, as long as we are aware of it.

Interactions of Factors and Marginal Selectivity

Let x and y be two processes in a directed, acyclic network. Suppose that experimental factor 1 affects process x and experimental factor 2 affects process y in the following way.

Let the duration of a process z when both factors are at their lowest levels be a random variable $D(z)$. A value taken by $D(z)$ will be denoted $d(z)$. Suppose when factor 1 is at level 2, the duration of process x is $D(x) + U$, where U is a non-negative random variable (see Footnote 2). Suppose the random variable corresponding to every process duration except that of x is unchanged when the level of factor 1 is changed. Suppose when factor 2 is at level 2, the duration of process y is $D(y) + V$, where $D(y)$ is as defined before and V is a nonnegative random variable. Suppose the random variable corresponding to the duration of every process except y is unchanged when the level of factor 2 is changed. That is, direct selective influence is assumed.

Let the random variable T_{ij} be the time to complete the task when factor 1 is at level i and factor 2 is at level j.

Now the statistic of concern in a 2×2 factorial experiment is the difference of mean differences or contrast in terms of the sample means $\bar{T}_{ij}$,

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

This corresponds to the second order theoretical difference of expected values

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

This important quantity will be the main focus of the sequel to this paper, because upon it rests the trichotomy method of classifying mental networks. Of course, if factors fail to act selectively in the intended manner then matters can go seriously awry. Even if the experimenter acts to directly affect the targeted processes, and in fact succeeds in this goal, stochastic dependencies in the network could cause an indirect nonselective influence as pointed out earlier. In order to forestall this from happening, certain conditions need to be imposed. These must be in the form of a further postulate on the distributions of the networks. The condition is weaker than the assertion of independence of the separate processes, along with certain distributional invariances, although that would be a special case of interest. Obviously if a system does not obey the postulate, then the predictions do not go through. Although dependencies are extremely important in processing networks, very little is known about how to identify them or show they may affect the results of psychological experiments. Some progress has recently been made by Colonius (1988) on this dimension for a certain class of experiments.

Definitions 7, 8, 9 and Theorem 1 state the positive side of selective influence. That is, we assume that direct selective influence via incrementing of durations occurs, and further, that no *direct* nonselective influence occurs. However, it is necessary to be explicit about what kinds of *indirect* nonselective influence are ruled out.

Our postulate that performs this duty appears to be the weakest required under the present circumstances. Except for T_{22}, the reaction times for the different conditions turn out to be linear combinations of $s(xr)$, $s(yr)$, $s(xy)$, t, u, v whenever they are nonzero. The random variable T_{22}, on the other hand, is a slightly more complex function (see sequel to this paper).

Because the various prolongation conditions can, in principle, lead to distinct marginal distributions, we affix a subscript (i, j) to the joint density function ($i = 1$ or 2 and similarly for j) indicating whether or not x and y have been prolonged respectively. Thus, $f_{12}(s(xr), s(yr), s(xy), v, t)$ indicates that y but not x has been prolonged and that this function may not in general (i.e., before imposition of the postulate given in Definition 10) equal

$$\int_{u=0}^{\infty} f_{22}(s(xr), s(yr), s(xy), u, v, t)\, du.$$

This would in turn mean that $E(V|(1, 2))$ might not equal $E(V|(2, 2))$, violating selective influence. Moreover, even if only V and not $D(y)$ were dependent on whether or not u was present, still untoward consequences erupt.

To put the latter verbally, these facts plus properties of expectations suggest that the weakest way to exclude indirect nonselective influence is to postulate that the densities for T_{ij} $(i, j = 1, 2)$ integrate to the same thing when we integrate over u or v.

A simple example of how matters can go awry in the absence of regularity conditions is found in serial processing with x followed by y, the only processes in the network. Here, $s(xy) = s(xr) = s(yr) = 0$ identically and $T_{11} = T$, $T_{21} = T + U$, $T_{12} = T + V$, $T_{22} = T + U + V$.

Let $D(x)$, $D(y)$ be distributed exponentially when unprolonged, but suppose that both $D(y)$ and V are conditionally dependent on $D(x)$ and U in a manner to be made explicit momentarily. Suppose also that when either is prolonged, an exponential random variable is added. We should like our theorems to produce additivity in $E(T_{ij})$ for serial systems but this may fail unless precautions are taken.

First, observe that even $T_{11} = T$ may be distributed differently in the various conditions, because $T = D(x) + D(y)$ and possibly $f(d(y)|d(x) + u) \neq f(d(y)|d(x))$

in general. Therefore, $E(T_{ij})$ would be expected to differ across conditions and would not cancel out when $\Delta E(T)$, the contrast, is computed.

Next we examine indirect selective influence of U on V in more detail. Let V depend on the overall duration of the first stage,

$$f_{12}(v \mid t) = f_{12}(v \mid d(x), d(y)) = f(v \mid d(x)) = e^{d(x)} b e^{-e^{d(x)} b v},$$

and under condition (2, 2), as in the following definition,

$$f(v \mid d(x), u) = e^{d(x)+u} b e^{-e^{d(x)+u} b v}.$$

In both cases V is stochastically speeded up when the first stage is long and vice versa, causing a negative dependence. By assumption, both $d(x)$ and u are distributed exponentially with, say, parameters a_1, a_2, respectively. Therefore, it follows that

$$E(V \mid (1, 2)) = \frac{1}{b} \cdot \frac{a_1}{1 + a_1}$$

$$E(V \mid (2, 2)) = \frac{a_1 a_2}{b} \left[\frac{1}{(a_2 - a_1)(a_1 + 1)} + \frac{1}{(a_1 - a_2)(a_2 + 1)} \right] = \frac{a_1 a_2}{b(a_1 + 1)(a_2 + 1)}.$$

Because these two expressions differ, they will not cancel out in $E(T)$, the contrast. The effect is to disturb the otherwise generically serial behavior of the network.

This example and others like it are precluded by postulating the following definition.

Definition 10. In 2×2 factorial experiment with direct selective influence at the incrementing level, let

1. $f_{22}(s(xr), s(yr), s(xy'), u, v, t)$
   = density for condition (2, 2) (both x, y incremented),
2. $f_{12}(s(xr), s(yr), s(xy'), v, t)$
   = density for condition (1, 2) (only y incremented),
3. $f_{21}(s(xr), s(yr), s(xy'), u, t)$
   = density for condition (2, 1) (only x incremented),
4. $f_{11}(s(xr), s(yr), s(xy'), t)$
   = density for condition (1, 1) (neither x, y incremented).

Then *marginal selectivity* holds if and only if

5. a. $\int_0^{\infty} f_{22}(s(xr), s(yr), s(xy'), u, v, t)\, du = f_{12}(s(xr), s(yr), s(xy'), v, t),$

b. $\int_0^\infty f_{22}(s(xr), s(yr), s(xy'), u, v, t)\, dv = f_{21}(s(xr), s(yr), s(xy'), u, t),$

c. $\int_0^\infty \int_0^\infty f_{22}(s(x, r), s(yr), s(xy'), u, v, t)\, du\, dv = f_{11}(s(xr), s(yr), s(xy'), t),$

6. $\int_0^\infty f_{12}(s(xr), s(yr), s(xy'), v, t)\, dv = f_{11}(s(xr), s(yr), s(xy'), t),$

7. $\int_0^\infty f_{21}(s(xr), s(yr), s(xy'), u, t)\, du = f_{11}(s(xr), s(yr)\, s(xy'), t).$

Note that 5. c is superfluous in the presence of parts 2 and 3 but is included for the reader's convenience. This definition therefore ensures that (1) The density f_{22} integrates over u to f_{12}, over v to f_{21}, and over both u, v to f_{11}. (2) The density f_{21} integrates over u to f_{11}. (3) The density f_{12} integrates over v to f_{11}. A useful outcome of Definition 10 for the succeeding paper is that one may start, say, with f_{11} and augment it with a dummy v to reach f_{12} and so on.

The reader can verify that the earlier serial example does not obey the foregoing conditions of marginal selectivity. Obviously marginal selectivity trivially precludes contamination by way of indirect nonselective influence. Moreover, there are two special cases of interest that fulfill marginal selectivity. First, we define *density invariance* to hold whenever two or more joint density functions of the same variables appearing in the various experimental (i, j) conditions are identical.

THEOREM 2. A. *If u and v are conditionally independent of one another, accompanied by invariance of densities, then marginal selectivity is satisfied. That is,*

1. $f_{22}(s(xr), s(yr), s(xy'), u, v, t) = f(s(xr), s(yr), s(xy'), t) \cdot f(u \mid s(xr), s(yr), s(xy'), t) \cdot f(v \mid s(xr), s(yr), s(xy'), t);$
2. $f_{12}(s(xr), s(yr), s(xy'), v, t) = f(s(xr), s(yr), s(xy'), t) \cdot f(v \mid s(xr), s(yr), s(xy'), t);$
3. $f_{21}(s(xr), s(yr), s(xy'), u, t) = f(s(xr), s(yr), s(xy'), t) \cdot f(u \mid s(xr), s(yr), s(xy'), t);$
4. $f_{11}(s(xr), s(yr), s(xy'), t) = f(s(xr), s(yr), s(xy'), t).$

B. *If all processes in the network are independent of one another and direct selective influence by incrementing on $d(x)$ and $d(y)$ are the only influences across the (i, j) conditions, then marginal selectivity is satisfied.*

Proof. Part A follows simply from integrating to the marginals. If B holds, then $s(xr)$, $s(yr)$, $s(xy')$, and t are all unaffected by prolongation of $d(x)$ or $d(y)$ or both. Furthermore, because u and v are 'placed' in distinct processes, they cannot interact with one another and neither affects the other random variables. These constraints permit the decomposition of part A. Q.E.D.

Observe that the conditions of A, B of Theorem 2 are definitely stronger than the imposition of marginal selectivity per se. The latter only requires that after integration, the various joint densities f_{ij} somehow turn out to be invariant. Theorem 2 gives strong but psychologically interesting sufficient conditions for this to be so.

What happens when factorial selectivity fails? The fact that the theorems showing the possibility of network identification do not go through does not in itself indicate whether matters are dire. There are no general theorems available showing specifically how absence of factorial selectivity could make, say, a parallel system look serial and so on. However, hints appear in Townsend and Ashby (1983, pp. 367–371 & 374–378) and in Townsend (1984, pp. 371–372 & 376–378). Also, Example 1 given earlier showed how the failure of selectivity through a stochastic dependency across processes could lead to nonadditivity in a serial system, contrary to expectation. Of course, *direct* nonselective influence could also cause problems. Nevertheless, investigation of how different types of nonselective influence can make systems mimic each other may be a first step in developing tests of, say, classes of the selective influence networks against possible mimicking nonselective influence networks. This will be the subject of a future paper.

A more global problem is the question of whether any of the postulates underpinning the present networks are testable. We imagine that few investigators would oppose the stochastic assumption. As stated earlier, the exhaustive processing at each node is relatively innocuous although it is of interest to develop theory/methodologies for, say, self-terminating networks or mixtures of exhaustive and other types. The discrete aspect of the processing postulates is definitely of psychological interest. We have little to add here about that issue but refer the reader to papers by Miller (1988), Proctor and Reeve (1985), Meyer *et al.* (1985), and Schweickert (in press). The partial ordering assumption seems sufficiently general that it is not at all obvious how one would go about testing it. To be sure, it would be nice to include the possibility of feedback loops, the main systemic lacuna of partially ordered networks. Finally, the selective influence postulate, as observed above, is critical and experimental tests would be of considerable strategic value. The employment of an $(n \times m)$ factorial experiment would check whether selective influence holds. As we have previously shown (Schweickert, Townsend & Fisher, 1987), the contrast sign (plus or minus) should be invariant over levels of the factor because it should depend only on the network structure. This result will be reviewed in the succeeding paper. Such an experiment should be useful to confirm or falsify our overall theory, including the postulate that the actual psychological system is a stochastic discrete mental network (i.e., the graph specifies PERT structure, etc.). If the sign changes in a strong manner, then any of the postulates (except stochasticity) may be at fault. In our opinion, empirical tests of selective influence will likely be tied closely to separation of selective from nonselective systems.

The requisite theoretical machinery is now assembled and the major classification theorems can proceed in the next paper.

REFERENCES

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

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

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

BORING, E. G. (1957). *A history of experimental psychology*. New York: Appleton-Century-Crofts.

COLONIUS, H., in press. Possibly dependent probability summation of reaction time. *Journal of Mathematical Psychology*.

DONDERS, F. C. (1868). Die Schnelligkeit Psychischer Processe. *Archiv für Anatomie und Physiologie*, 657–681. [On the speed of mental processes. In W. G. Koster (Ed. and Trans.), *Attention and performance II*. Amsterdam: North-Holland. Reprinted from *Acta Psychologica*, 1969, **30**.]

EGETH, H., JONIDES, J., & WALL, S. (1972). Parallel processing of multi-element displays. *Cognitive Psychology*, 3, 674–698.

ERIKSEN, C. W., & HOFFMAN, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, **12**, 201–204.

FISHER, D. L. (1982). Limited-channel models of automatic detection: Capacity and scanning in visual search. *Psychological Review*, **89**, 662–692.

FISHER, D. L. (1984). Central capacity limits in consistent mapping, visual search tasks: Four channels or more? *Cognitive Psychology*, **15**, 449–484.

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

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

GOLDSTEIN, W. M., & FISHER, D. L., in press. Stochastic networks as models of cognition: Derivation of response time distributions using the order-of-processing method. *Journal of Mathematical Psychology*.

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

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.

KOHFELD, D. L., Santee, J. L., & WALLACE, N. D. (1981a). Loudness and reaction time, I. *Perception & Psychophysics*, **29**, 535–549.

KOHFELD, D. L., SANTEE, J. L., & WALLACE, N. D. (1981b). Loudness and reaction time. II. Identification of detection components at different intensities and frequencies. *Perception & Psychophysics*, **29**, 550–562.

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

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

MCGILL, W. J. (1963). Stochastic latency mechanisms. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical psychology* (Vol. I). New York: Wiley.

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

MEYER, D. E., & SCHVANEVELDT, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90 (2), 227–234.

MEYER, D. E., YANTIS, S., OSMAN, A. M., & SMITH, J. E. K. (1985). Temporal properties of human information processing: Tests of discrete vs. continuous models. *Cognitive Psychology*, **17**, 445–518.

MILLER, J. O. (1982a). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 273–296.

MILLER, J. O. (1982b). Effects of noise letters on decisions : discrete or continuous flow of information? *Perception & Psychophysics*, 31, 227–236.

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

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*. Hillsdale, NJ: Erlbaum.

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

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

PROCTOR, R. W., & REEVE, T. G. (1985). The compatibility effects in the assignment of symbolic stimuli to discrete finger responses. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 623–639.

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

SCHWEICKERT, R. (1977). A generalization of the additive factor method. (Tech. Rep. No. 77–2). Ann Arbor: University of Michigan, Mathematical Psychology Program.

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

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

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

SCHWEICKERT, R. (1983a). Synthesizing partial orders given comparability information: Partitive sets and slack in critical path networks. *Journal of Mathematical Psychology*, 27, 261–276.

SCHWEICKERT, R. (1983b). 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. (in press). Separable effects of factors on activation functions: The rectangle condition applied to d' and evoked potentials. *Psychological Bulletin*.

SCHWEICKERT, R., TOWNSEND, J. T., & FISHER, D. L. (1987). *Simple tests distinguishing sequential and concurrent processes in task networks*. Paper presented at Twenty-eight Annual Meeting of the Psychonomic Society, Seattle.

SHIFFRIN, R. M., & SCHNEIDER, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.

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

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

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

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

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

TOWNSEND, J. T. (1969, April). Mock parallel and serial models and experimental detection of these. Purdue Centennial Symposium on Information Processing, Purdue Press, Purdue University.

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

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

TOWNSEND, J. T. (1974a, August). *Independent parallel models and analyses of reaction time distributions.* Paper presented at Seventh Annual Mathematical Psychology Meetings, Ann Arbor, Michigan.

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

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

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

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

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

TOWNSEND, J. T. (1986). *Relationships of nonparametric aspects of probability distributions.* Nineteenth Annual Meeting of Society for Mathematical Psychology, Boston, Massachusetts.

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

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

TOWNSEND, J. T., & EVANS, R. (1983). A systems approach to parallel-serial testability and visual feature processing. In H. G. Geissler, H. F. J. M. Buffart, E. L. J. Leeuwenberg and V. Sarris (Eds.), *Modern Issues in Perception.* Berlin: VEB Deutsches Verlag der Wissenschaft.

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

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

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

VORBERG, D. (1988). Network models of reaction time. Paper presented at Twenty-fourth International Congress of Psychology. In S. W. Link and H. C. Micko (Conveners), *Response Latency: Theories of Psychological Processes.* Sydney, Australia.

WUNDT, W. (1896). *Grundriss der Psychologie.* Leipzig, Engelman.

RECEIVED: December 22, 1986

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