Journal of Mathematical Psychology **41**, 392–397 (1997)
Article No. MP971185

Parallel Processing Response Times and Experimental Determination of the Stopping Rule*

James T. Townsend

Indiana University

and

Hans Colonius

University of Oldenburg

It was formerly demonstrated that virtually all reasonable exhaustive serial models, and a more constrained set of exhaustive parallel models, cannot predict critical effects associated with self-terminating models. The present investigation greatly generalizes the parallel class of models covered by similar ''impossibility'' theorems. Specifically, we prove that if an exhaustive parallel model is not super capacity, and if targets are processed at least as fast as non-targets, then it cannot predict such (self-terminating) effects. Such effects are ubiquitous in the experimental literature, offering strong confirmation for self-terminating processing.

INTRODUCTION

Since the 1960's, a number of advances in the theory and related methodology of human information processing have taken place. Obviously much of this progress has been in the form of relatively specific models and theories (e.g., Burbeck & Luce, 1982; Egeth, 1966; Falmagne & Theios, 1969; Green & Luce, 1972; Link & Heath, 1975; Ratcliff, 1978; Sternberg, 1966). Among these and subsequent references, most have emphasized reaction time (RT) as the main dependent variable of interest.

However, a more meta-theoretic approach has also appeared wherein large classes of models are analyzed with regard to their predictions in specified types of experimental designs. Some of these have helped by indicating equivalence classes of models of opposing psychological principles and thus revealing where common experimental paradigms could not test such principles against one another. An example is the ubiquitous parallel (simultaneous) vs serial (one at a time) processing issue (e.g., Townsend, 1969, 1972).[1] Investigators have also probed such questions with the aid of model simulation (e.g., Ashby, 1982a; Ratcliff, 1988).

Other meta-theoretical studies have explored fundamental means of experimentally deciding such questions (e.g., Ashby, 1982b; Ashby & Townsend, 1980; Schweickert, 1978, 1983, 1985; Schweickert & Townsend, 1989, Schweickert & Wang, 1993; Sternberg, 1969; Townsend, 1976a,b; Townsend & Evans, 1983; Townsend & Schweickert, 1989). Computational advances on complex models of information processing have also transpired (e.g., Fisher & Goldstein, 1983; Goldstein & Fisher, 1991, 1992). Many of these investigations, of all three varieties, have been reviewed in Luce (1986), Townsend and Ashby (1983), and Townsend (1990a).

In addition to processing architecture, another problem of considerable import has been the stopping rule people's cognitive systems employ when sufficient information has been acquired to make a correct response. For instance, when search of short-term memory may occupy a few hundred milliseconds, it may be more efficient to simply process all items "exhaustively" rather than cease when a positive match is located (e.g., see Sternberg, 1966). On the other hand, if it can be done efficiently, why process useless information after a target has been located, that is, why not simply "self-terminate"?

Sternberg's (1966) short-term memory search and the Atkinson, Holmgren and Juola (1969) rapid visual display search experiments contained trials where no targets existed and trials where exactly one positive target was placed. The emphasis in an immense number of experiments since that time has remained on those and similar paradigms

* The authors' collaboration was supported by a grant from the American Council of Learned Societies and the Deutsche Akademische Austauschdiest and National Science Foundation Grant BNS # 9112813.

[1] However, even early papers on the parallel-serial issue were not totally concentrated on the "negative" message of model mimicking. For instance, Townsend (1972) begins to look for, and suggest ways that might more fundamentally test these issues.

0022-2496/97 $25.00

(Van Zandt & Townsend, 1993), although other designs, such as where every stimulus in the search set is a target, offer valuable information on architecture and stopping rule (Egeth & Mordkoff, 1991).

There have been two predominant observable variables used to test self-termination vs. exhaustive processing over the past quarter century. The first is mean *RT* slope as a function of search set size ($n =$ load). Consider the prototypical "standard serial" model where processing takes place on items sequentially, without temporal overlap, and where the average processing time on the items is invariant over all aspects of the experiment (e.g., processing position, stimulus location, positive vs. negative match and number of items in the search set). When processing in this type of model is self-terminating, it predicts slopes on positive trials that are one-half those on negative trials. On the other hand, when processing is exhaustive, the slopes of the positive and negative functions of load must be equal in the standard serial model.

The other main-stay dependent variable in the stopping rule debate is mean *RT* as a function of position of the target (on positive trials of course) in the stimulus set. The idea is that in self-terminating serial models, the position may affect *RT* because a non-equality in frequency of different processing paths through the items would imply that some positions would typically finish before others. In exhaustive (standard) serial models, the position should not have any effect because all positions have to be completed on every trial.

Parallel self-terminating models can readily predict the same type of slope relationships as the standard serial model and parallel exhaustive models can readily predict the same relationships as standard serial exhaustive models. Similarly, non-standard serial self-terminating models can easily make predictions characteristic of common parallel self-terminating models and the same goes for non-standard serial exhaustive models with regard to parallel exhaustive processing (for these and other predictions of mean *RT* as a function of load, see Townsend, 1974, or the summary in Townsend & Ashby, 1983, Chapter 4). Thus, the stopping problem crops up in both parallel and serial architectures. This establishes the background when stopping rule is the same, but architecture varies. What happens when one attempts to produce the effects commonly predicted by one stopping rule (e.g., self-termination) with a model based on the opposite stopping rule (exhaustive)?

It was noticed early-on that standard serial self-terminating models could handily predict position functions associated with exhaustive models when all processing paths were equally likely (e.g., Atkinson, Holmgren & Juola, 1969). Similarly, parallel models with everywhere equal processing speeds could do the same.

Furthermore, when non-standard serial and a variety of parallel models began to be investigated, it became apparent that relaxing assumptions about processing speed (more exactly, about the distribution of processing times) on the items could lead to the ability of self-terminating models to predict the standard exhaustive result of equal slopes of positive and negative load functions. Thus, it was discovered relatively soon that self-terminating models of either variety (parallel or serial) could encompass exhaustive predictions with little difficulty, and in an intuitive manner (summarized in Townsend, 1974; Townsend & Ashby, 1983, Chapter 7 and, Townsend & Roos, 1973). The upshot was that self-terminating models were proven to comfortably accommodate the predictions typifying exhaustive processing.

Even more dispiriting, from the model testing point of view, was the theoretical demonstration that (nonstandard) serial or parallel exhaustive models were capable of predicting a certain degree of positive vs. negative slope deviations and even of position effects (Ashby, 1976; Townsend, 1974). The picture began to look as bleak as with the parallel-serial issue with regard to a model mimicking impasse. That is, exhaustive models gave the appearance of being able to mimic self-terminating effects to some extent. The case seemed to be rapidly converging toward a conclusion that slope and position data were useless in terms of identifying type of stopping rule in search experiments.

It was therefore a pleasant surprise when Townsend and Van Zandt later (1990) showed that although a very general class of serial exhaustive models (e.g., much larger than the class of standard serial models) could indeed predict some slope differences and position effects, they were *severely limited* in their ability to do so. Similar theorems for certain parallel classes of models were proffered. These findings were followed up by further theoretical results and a literature review (Van Zandt & Townsend, 1993). Hence, the general conclusion could now be advanced that there exists an asymmetry in possible conclusions from slope and position experimental data: Data seemingly in favor of exhaustive processing could readily be produced by self-terminating models (parallel or serial), but data supporting self-termination falsified exhaustive models of processing.

Nevertheless, a qualification remained in that the class of parallel models for which comparable results were accomplished was significantly less general than that of the serial models. The purpose of this paper is to greatly broaden the class of parallel models covered by impossibility theorems (i.e., of producing certain types of position and slope effects).

THEORETICAL RESULTS

The critical experimental devices discussed above, negative-to-positive mean *RT* slope ratio and positive trial mean RT position curve, are intimately related to the notion of processing capacity. Roughly, the more capacity available to a processing mechanism, the faster it can perform its function aid vice versa. "Unlimited capacity," again roughly, implies that speed of processing is unaffected by an increase in processing load, the latter often measured by the number of items or tasks to be processed. "Limited capacity" means speed is slowed whereas "super capacity" implies the surprising result of an actual improvement in processing speed as the load increases.

We now develop a more precise notion of capacity. Consider any event which possesses a finishing time, associated with a set of items being processed. The system is said to be unlimited capacity with respect to that event if the marginal distribution function on that event is unaffected by the number of items, n, in the processing set. If at every time t, the distribution function decreases as n increases, we have limited capacity and if it increases, we have super capacity.[2] This general format encompasses the fact that capacity can be "measured" at a relatively "micro" level such as the item level or even feature level, or at a more "macro" level such as the maximum or minimum processing time of all items (Townsend, 1974). Hence, it can be seen that capacity might be unlimited with regard to the individual item level, yet limited capacity at, say, the exhaustive processing level. This can, and often does, happen for purely stochastic reasons (Colonius & Vorberg, 1994; Townsend & Ashby, 1983). For instance, even with unlimited capacity independent parallel channels, the exhaustive processing time will stochastically increase due to statistical effects.

In Proposition 1, on the slope ratio, it is sufficient to state the conditions on capacity in terms of the average distribution function on exhaustive processing time, over the placement of the positive target. In Proposition 2, on strong position effects, it is necessary to fix the target's position and state the capacity condition. The capacity condition is still imposed on the maximum (i.e., exhaustive) processing times.

Basically, the conditions rule out super capacity models where the probability distribution functions increase (i.e., for all $t > 0$; a distributional way of indicating "speed-up") as a function of increasing n. They also rule out highly Gestalt sorts of processing where distribution function on exhaustive processing times can somehow stay quite large as n grows, despite the necessity of finishing all of the negative matches. Gestalt formulations would be one of the few rationales as to why capacity might improve with load. For instance, consider strings of letters such that for $k < n$, the strings are meaningless but at n, the string makes a word. Then capacity might appear to be super, relative to the specific materials, since RT would speed up at $k = n$.

[2] The level of precessing capacity, can, in fact, be assessed within probability distributions on processing time by distinct aspects of the distribution. A very coarse (and rather weak) index would be by the magnitude of mean RTs (smaller mean implies more capacity) up to the quite fine-grained (and quite strong) magnitude of hazard functions (see, e.g., Townsend, 1990b; Townsend & Ashby, 1983). Capacity assessed at the distribution ordering level, as in this paper, is at an intermediate level (e.g., it implies an ordering of means but is itself implied by an ordering of hazard functions).

Assume the targets are placed in each position with probability $1/n$ by the experimenter (easily generalized). Let $\mathbf{M}_n = \text{MAX}(\mathbf{T}_1, ..., \mathbf{T}_n)$, the joint (*exhaustive*) finishing time random variable's distribution, be written, when n channels are functioning as, **TARGET ABSENT**:

$$F_n(t) = P[\mathbf{T}_1^- \leqslant t, ..., \mathbf{T}_n^- \leqslant t]$$

with expectation:

$$E[\text{Max}(\mathbf{T}_1^-, ..., \mathbf{T}_n^-)] = E(\mathbf{M}_n).$$

SINGLE TARGET PRESENT (with positive target match in position i):

$$F_{n,i^+}(t) = P[\mathbf{T}_1^- \leqslant t, ..., \mathbf{T}_i^+ \leqslant t, ..., \mathbf{T}_n^- \leqslant t]$$

with expectation:

$$E[\text{Max}(\mathbf{T}_1^-, ..., \mathbf{T}_i^+, ..., \mathbf{T}_n^-)] = E(\mathbf{M}_{n,i^+}).$$

The $\mathbf{T}_i^-$ random variable in the target-absent expression bears no necessary relation to $\mathbf{T}_i^-$ appearing in the target-present expression, beyond those conditions appearing in the propositions.

Now consider two straight lines for mean RT, the first for positive $RT^+(n) = an + b$ and the second for negative $RT^-(n) = cn + d$. For instance, in a standard serial model with self-termination, $a = \frac{1}{2}$ and $c = 1$. The result is $c/a = 2$ and in general we investigate the inequality $c/a \geqslant 2$. It is easy to see that this inequality implies a 2-step inequality: $RT(n+1) - RT(n) \geqslant RT^+(n+2) - RT^+(n)$. However, we wish to encompass nonlinear mean RT curves with the possibility of distinct processing speeds on distinct items and positions, including faster "rates" for positive matches. We therefore express our generalized 2-step inequality as:

Generalization of 2:1 *Slope Ratio Prediction*:

"Positive" mean RT (2-step) Difference

$$\begin{aligned} &= \frac{1}{n+2} \sum_{i=1}^{n+2} E(\mathbf{M}_{n+2,i^+}) - \frac{1}{n} \sum_{i=1}^{n} E(\mathbf{M}_{n,i^+}) \\ &\leqslant E(\mathbf{M}_{n+1}) - E(\mathbf{M}_n) \\ &= \text{"Negative" Mean } RT \text{ (1-step) Difference} \qquad (1) \end{aligned}$$

Standard serial self-terminating models plus a very large class of other self-terminating parallel and serial models can predict this inequality. Proposition 1 demonstrates that most parallel exhaustive models predict violations for any value of n.

PROPOSITION 1. *Assume, for each n,*

$$\frac{1}{n+1}\sum_{i=1}^{n+1} F_{n+1,i^+}(t) < F_n(t) < \frac{1}{n}\sum_{i=1}^{n} F_{n,i^+}(t)$$

i.e., (i) *positive processing* (*averaged over position*) *is faster than negative processing* (*the upper bound*) *and* (ii) *capacity is sufficiently limited* (*the lower bound*). *Then the* Generalized 2:1 Slope Prediction is impossible.

Proof. Replacing n by $n+1$ in the lower inequality of Proposition 1 gives

$$\frac{1}{n+2}\sum_{i=1}^{n+2} F_{n+2,i^+}(t) < F_{n+1}(t);$$

Adding $F_n(t)$ on the left side and adding the (larger) $(1/n)\sum_{i=1}^{n} F_{n,i^+}(t)$ on the right side yields

$$\frac{1}{n+2}\sum_{i=1}^{n+2} F_{n+2,i^+}(t) + F_n(t) < \frac{1}{n}\sum_{i=1}^{n} F_{n,i^+}(t) + F_{n+1}(t).$$

Rewriting this in terms of survivor functions and integrating from 0 to ∞ reverses the inequality of the slope ratio prediction. Q.E.D.

As with a prototypical parallel model meeting the conditions of Proposition 1, assume stochastic independence. Also, assume that each negative distribution on separate positions is just $P_n[\mathbf{T}_i^- \leqslant t] = G^-(t, n)$ across all positions. Likewise, the positive distribution functions are $P_n(\mathbf{T}_i^+ \leqslant t) = G^+(t, n) > G^-(t, n)$. Assume too that $G^-(t, n) > G^-(t, n+1)$. Then the condition of Proposition 1 becomes

$$[G^-(t, n+1)]^n\, G^+(t, n+1) < [G^-(t, n)]^n < [G^-(t, n)]^{n-1}\, G^+(t, n).$$

and is clearly satisfied. As a further example, retaining stochastic independence but allowing for different distribution functions for the different positions, one gets for all n and i,

$$F_n(t) = \prod_{j=1}^{n} G_{n,j}^-(t)$$

and

$$F_{n,i^+}(t) = \prod_{\substack{j=1\\ j\neq i}}^{n} G_{n,j}^-(t)\, G_{n,i}^+(t).$$

The conditions of Proposition 1 are then satisfied if, for all t and n.

$$G_{n,i}^-(t) < G_{n,i}^+(t) < G_{n-1,i}^-(t).$$

In order to motivate the strong position effect notion, observe that in self-terminating *serial* processing with a single processing path, processing stops when the target in position i is reached. Now recall that in standard serial processing, and on positive data, this implies $E(\mathbf{RT}) = E[\sum_{j=1}^{i} \mathbf{T}_j] = iE(\mathbf{T})$. This fact in turn implies that the self-terminating stopping time, for fixed position i (containing the target) is invariant across set size ($=$ load $= n$) for $n \geqslant i$. Consider a plot of mean response time as a function of target position i. Then a consequence of the above equation for mean response times is that as n increases, each successive position curve lies exactly on top of the one for $n-1$, except one extra point is added for $i = n$. It should be evident that the standard exhaustive model predicts the processing time for any position to be $nE(\mathbf{T})$ and hence violates the strong position prediction. More general serial models also fail to predict strong position effects (Townsend & Van Zandt, 1990). On the other hand, it is easy to show that self-terminating *parallel* models can readily produce strong position effects (e.g., Townsend, 1974; Townsend & Van Zandt, 1990; Van Zandt & Townsend, 1993). What about exhaustive parallel models?

In our present generalized version of strong position effects, we simply request that the response time for any given position i, be invariant across all set sizes greater than or equal to i. Thus, the position functions need not be linear. We shall show that again, a very large class of exhaustive parallel models, which are not super capacity, are incapable of predicting these effects.

Generalization of Strong Position Effects

For any $i = 1, 2, \ldots$ and any $i \leqslant n_1 \leqslant n_2$,

$$\begin{aligned} &E[\mathrm{Max}(\mathbf{T}_1^-, \ldots, \mathbf{T}_i^+, \ldots, \mathbf{T}_{n1}^-)] = E(\mathbf{M}_{n1,i^+}) \\ &\quad = E[\mathrm{Max}(\mathbf{T}_1^-, \ldots, \mathbf{T}_i^+, \ldots, \mathbf{T}_{n2}^-)] = E(\mathbf{M}_{n2,i^+}). \end{aligned} \tag{2}$$

PROPOSITION 2. *Assume that on positive trials, capacity is limited on the diagonal in the sense that for any* $i = 1, 2, \ldots n_1 < n_2$,

$$\begin{aligned} &P(\mathbf{T}_1^- \leqslant t, \ldots, \mathbf{T}_i^+ \leqslant t, \ldots, \mathbf{T}_{n1}^- \leqslant t) \\ &\quad = F_{n1,i^+}(t) > F_{n2,i^+}(t) \\ &\quad = P(\mathbf{T}_1^- \leqslant t, \ldots, \mathbf{T}_i^+ \leqslant t, \ldots, \mathbf{T}_{n2}^- \leqslant t). \end{aligned} \tag{3}$$

Then strong position effects are impossible.

Proof. By the above condition (Eq. (3)), the two sides of Eq. (2) become

$$E[\mathbf{M}_{n1,\,i^+}) = \int_{t=0}^{\infty} [1 - F_{n1,\,i^+}(t)]\, dt$$
$$< E[\mathbf{M}_{n2,\,i^+}) = \int_{t=0}^{\infty} [1 - F_{n2,\,i^+}(t)]\, dt, \quad (4)$$

in contradiction to Eq. (2), and the proposition is proven. Q.E.D.

Note that capacity is limited with regard to exhaustive processing (i.e., on the diagonal as above) even if it is unlimited capacity at the individual item level.

We again examine the class of models similar to the first example of Proposition 1:

$$\begin{aligned} F_{n1,\,i^+}(t) &= [G^-(t, n_1)]^{n_1 - 1}\, G^+(t, n_1) \\ &> [G^-(t, n_2)]^{n_2 - 1}\, G^+(t, n_2) \\ &= F_{n2,\,i^+}(t), \qquad n_1 < n_2 \end{aligned} \quad (5)$$

Equation (5) is satisfied as long as

$$G^-(t, n_1) > G^-(t, n_2) \qquad \text{and} \qquad G^+(t, n_1) > G^+(t, n_2),$$

the usual condition for limited-to-unlimited capacity, on negative or positive matches, respectively.

DISCUSSION

These results reinforce those of Townsend and colleagues (Ashby, 1976; Townsend, 1974; Townsend & Roos, 1973; Townsend & Van Zandt, 1990; Van Zandt & Townsend, 1993). Vast classes of parallel and serial models, when assumed to be exhaustive, are mathematically incapable of producing strong position effects, or 2:1 or greater (negative to positive) slope ratios. On the other hand, it was formerly demonstrated that self-terminating models of either ilk, can readily encompass both those features, as well as predictions usually associated with standard serial or parallel exhaustive processing.

The conditions for the propositions seem quite reasonable. It is true that some of our and Townsend and Van Zandt's (1990) results depend on positive matches being faster than negative. However, data where negative match *RT*s are both faster and yet evidence a slope that is twice that of the positive match functions are exceedingly rare (in light of the enormous number of search experiments carried out over the last few decades we hesitate to claim there are none, but we are unaware of any). It is possible in principle that negative match *RT* curves generally lie above the positive curves due to mechanisms other than the comparison process, so that faster positive processing times could accompany overall slower *RT*s but this seems to be a bit tortured in light of the preponderance of evidence for fast positive matches. Nevertheless, it might be of some interest to try to generalize the present and other findings where faster negative processing times are a reasonable possibility.

When the collated theoretical findings are compared with the plethora of empirical results in such paradigms over the past quarter century, we believe it must be concluded that self-termination is a fact. Indeed, examples of highly general (in terms of numbers of parameters) exhaustive parallel models tested against typical data, provided concrete demonstration of these claims (Van Zandt & Townsend, 1993). Further, although data suggestive of exhaustive processing are sometimes found, self-terminating models could also be behind those results. Thus, if parsimony is brought to bear, self-terminating models appear to be the preferred choice.

REFERENCES

Ashby, F. G. (1976). Pattern matching: Self-terminating or exhaustive processing? Unpublished masters thesis, Purdue University.

Ashby, F. G. (1982a). Deriving exact predictions from the cascade model. *Psychological Review*, **89**, 599–607.

Ashby, F. G. (1982b). 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*, **21** (2), 93–123.

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

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

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

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

Egeth, H. E., & Mordkoff, T. J. (1991). Redundancy gain revisited: Evidence for parallel processing of separable dimensions. In G. R. Lochead and J. R. Pomerantz (Eds.), *The perception of structure*. Washington, D.C.: American Psychological Association.

Falmagne, J. C., & Theios, J. (1969). On attention and memory in reaction time experiments. *Acta Psychologica*, **30**, 316–323.

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.

Goldstein, W. M., & Fisher, D. L. (1991). Stochastic networks as models of cogntion: 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** (1), 129–145.

Green, D. M., & Luce, R. D. (1972). Speed-accuracy tradeoff in auditory detection. In S. Kornblum (Ed.), *Attention and performance* (Vol. 4). New York: Academic Press.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, **40**, 77–105.

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

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59–108.

Ratcliff, R. (1988). A note on mimicking additive reaction time models. *Journal of Mathematical Psychology*, **32**, 192–204.

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

Schweickert, R. (1983). Latent network theory: Scheduling of processes in sentence verification and the stroop effect. *Journal of Experimental Psychology*: *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., & 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.

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.

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

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. Hillsdale, NJ: Lawrence Erlbaum.

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

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

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

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

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.

Townsend, J. T., & Evans, R. (1983). A systems approach to parallel-serial testability and visual feature processing. In H. G. Geissler (Ed.), *Modern issues in perception*, pp. 166–189. Berlin: VEB Deutscher Verlag der Wissenschaften.

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. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, **33**, 309–327.

Townsend, J. T., & Van Zandt, T. (1990). New theoretical results on testing self-terminating vs. exhaustive processing in rapid search experiments. In H. G. Geissler (Ed.), *Psychological explorations of mental structures*, pp. 469–489. Toronto: Hogrefe & Huber.

Van Zandt, T., & Townsend, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual and memory search: An evaluative review. *Perception* & *Psychophysics*, **53** (5), 563–580.

Received: August 7, 1997