

A clarification of self-terminating versus exhaustive variances in serial and parallel models

JAMES T. TOWNSEND

Indiana University, Bloomington, Indiana

A recent study published in Perception & Psychophysics (Donnelly, Found, & Müller, 1999) employs response time (RT) variances (in the form of standard deviations) in addition to mean RTs. Variances can contribute greatly to model testing. However, there is a danger of perpetuating the kinds of logical and methodological errors that have long attended research employing mean RTs alone. This commentary clarifies the theoretical and methodological issues, points out some new results concerning variability in search processes, and indicates how to resolve the global and specific challenges associated with identifying psychological mechanisms.

This note offers a commentary accompanied by the clarification of quantitative demonstrations regarding elementary architectures and stopping rules. We are concerned only with the prototypically elementary architectures of *serial* (one-at-a-time) versus *parallel* (simultaneous) type. *Stopping rule* refers to the issue of whether search can cease on trials when a target is present as soon as it is located. These issues will be further elaborated below. It has been observed that the theoretical obstacles associated with experimental determination of architecture and stopping rule are to be found in much inference regarding psychological mechanisms in the literature. The importance of mathematical research in establishing where models can and cannot be distinguished is underlined by a contrast of the situation with architecture versus that with stopping rule. In the situation of architecture, quantitative results showed that the method of varying the load (search set size, n) and measuring response time (RT) does not by itself lead to clear empirical distinctions of serial versus parallel processing. Mathematical work has further substantiated alternative methods of attaining this end. On the other hand, comparable quantitative results have revealed that the varying of n (i.e., the same experimental variation that is weak regarding architecture) and the study of such dependent variables as RT slopes and position effects does strongly discriminate stopping rules. The latter situation will be discussed later in the paper.

Donnelly, Found, and Müller (1999), in an interesting study of visual search for impossible objects, are in danger of extending logical errors typically made with re-

gard to mean RTs to RT variances. In five experiments, they investigated the extent to which form (shape) and metric attributes (three-dimensional orientation) are processed in early vision. Both types of attributes were defined by relations between line elements. The time honored tack of varying the number of items in the visual display and recording the RTs was implemented (e.g., Atkinson, Holmgren, & Juola, 1969; Egeth, 1966). It was concluded from the pattern of RTs that search for an abstract property, in this case, whether the form constituted a real or impossible object, was slow and serial. In contrast, the authors argued from their results that search for orientation (see, e.g., Enns & Rensink, 1990) was parallel, although probably of limited capacity.

This note is in no way an attack on that particular paper. In fact, these authors (i.e., Donnelly et al., 1999) are aware of the difficulties in distinguishing parallel versus serial models, and the general theoretical and experimental tactics appear generally suitable to this author. There were also a number of aspects of the study unrelated to the present issue. Nevertheless, some of the central conclusions were affected by this problem, and, since this type of predicament continues to appear in the experimental literature, it seems worthwhile to clarify the situation in the present specific instance. This note attempts to constitute more than an admonition, since some new results and techniques pertinent to the use of variances of RT are brought to light. Also, the specific message is part of a general caution that is relevant to psychological research in general.

One of the logical causes of difficulty in the present investigation concerns a wide-spread, but not inevitable obstacle, in psychological research. That obstacle is the inclination to assess the predictions of a particular model or theory that possesses a certain property and then to proceed to act as though every model that retains that property makes the same predictions. Over the past several decades, there has been a mountain of research on parallel and serial processing that makes that logical error, despite (as Donnelly et al., 1999, note with regard to the slopes of mean RT), repeated attempts to correct it (e.g., Atkinson et al., 1969; Murdock, 1971; Townsend, 1971, 1990b; Townsend & Ashby, 1983, Chapters 4, 6, and 7). Unfortunately, the mistake does not become less harmful to scientific progress simply because it continues to be made across generations.

Let us be clear that, if all serial models predicted increasing linear RT functions and no parallel models could make that prediction, using that characteristic to test parallel versus serial processing would provide correct inferences. This is not the case: Quite specific predictions that distinguish only very special cases of parallel versus serial models are used over and over by investigators to claim the generic serial or parallel results—this is simply

Correspondence concerning this article should be addressed to J. T. Townsend, Indiana University, Department of Psychology, 1101 E. 10th St., Bloomington, IN 47405-7007 (e-mail: jtownden@indiana.edu).

wrong. It would be logically correct, if not quite so interesting, to make the restricted claim that the very specific models (only) were distinguished, but this is virtually never done. Moreover, although this kind of error is perhaps most salient in the arena of simple mental architectures, it appears to be quite ubiquitous in psychological research in general.

Sternberg's (1966) seminal study concluded that increasing, relatively straight line mean RT functions (of n) are compatible with serial processing but not parallel processing. However, the falsified parallel models were confined to what we would now call the class of unlimited capacity, independent parallel models. Furthermore, almost all subsequent studies have leaped to the much stronger and unfounded deduction that increasing straight lines imply serial processing and nothing else.

An even more egregious inference has been that only flat RT functions of n are predicted by parallel models. Even in his 1966 study, Sternberg recognized that with an exhaustive (all items processed) stopping rule, parallel models whose individual processing times possessed nonzero variances typically would not predict flat RT functions. Furthermore, the mimicking pitfall is not at all confined to those particular examples (i.e., straight line predictions by serial, flat functions by parallel) of serial or parallel models. In fact, it has been known for many years that, in the absence of specially constructed paradigms, exact mimicking of serial models by parallel models of any variety, and vice versa, pose a significant dilemma (e.g., Atkinson et al., 1969; Luce, 1986; Murdock, 1971; Townsend, 1969; 1976a, 1976b; Vorberg, 1977). In addition, a rule-of-thumb criterion for seriality, mean RT slopes of greater than 10 msec, has been shown to be seriously problematic (Wolfe, 1998).

Among promising approaches for experimental determination of architecture (Townsend, 1990b) are those based on factorial strategies (e.g., for factorial approaches, see Schweickert, 1978; Schweickert & Townsend, 1989; Sternberg, 1998; Townsend, 1984; for other tests see, e.g., Ashby, Tein, & Balakrishnan, 1993; Neufeld & McCarty, 1994; Snodgrass & Townsend, 1980; Townsend, 1976b). Interestingly, the magnitude of mean interaction contrast, a key statistic for discriminating serial from parallel processes, is sometimes a function of the individual item processing time variances (Egeth & Dagenbach, 1991; Townsend & Diederich, 1996).

A valid strategy in attempting to narrow the body of explanatory theory for natural phenomena is to increase the set of dependent variables under scrutiny. In the present study, the authors take the admirable step of bringing in standard deviations of RT, rather than only means, to increase the weight of evidence, hopefully converging on a conclusion. Nevertheless, it is critical not to carry the old logical error into the new milieu. Even though variances or standard deviations can help to reduce the set of candidate models, the general mimicking theorems showing equivalence between large classes of serial and parallel models immediately imply that any statistic (mean, standard deviation, etc.) will be identical for mutually

mimicking models. We now turn to considering trouble spots for variance (equivalently, standard deviation) predictions with regard to architecture and, in particular, in relation to the Donnelly et al. (1999) study, discuss some interesting facets of variances in standard kinds of parallel models and exhibit ways of setting up equations to study the behavior of variances in parallel models. Although the stopping rule concept played a role in the target study on architecture, its determination was not the main goal. A final section outlines the more positive conclusions stemming from mathematical results regarding the stopping rule issue when employing the traditional RT measures and varying the load n , in contrast to the problematical situation for architecture.

Response Time Variances and Architecture

Donnelly et al. (1999) found relatively sizeable mean RT slopes as display size varied, with both impossible and possible figures, although the former were, to be sure, of truly impressive magnitude. In certain conditions, there were target-absent versus target-present mean RT slope differences that were related by the authors to stopping rule. As noted earlier, the stopping rule refers to the issue of whether search can cease when a target is located (*self-termination*) or always continues until all the items in the search set have been checked (*exhaustive processing*). In experiments with high accuracy, it is assumed that target-absent trials are always accompanied by exhaustive processing since all the items must be checked before a correct response can be made. The authors realized that, although the striking difference between search slopes for impossible versus possible figures could be due to distinct architectures (serial vs. parallel), in principle, it could be caused by differences in capacity limitations per se. In fact, a parallel model, the sum of whose processing rates is always the same regardless of load (i.e., fixed capacity over display size here) and across all stages of processing (due to reallocation of capacity of finished items to uncompleted items), makes identical predictions to the serial model (e.g., Townsend, 1971).

Thus, they (Donnelly et al., 1999) turned to potential predictive dissimilarities for standard deviations to discriminate parallel versus serial processing. In particular, they discovered that, although standard deviations increased more on target-present trials than on target-absent trials in the impossible searches, target-present or target-absent standard deviations increased about the same amount across display size on impossible figure searches (e.g., pp. 678-679). The subsequent inference was that search for orientation could be limited capacity parallel, whereas search for impossible figures takes place "certainly via a serial self-terminating search" (p. 689). I would agree that the search for impossible figures is very limited capacity, but the deductions regarding serial versus parallel processing should be adjudged premature.

The conclusions in the target paper (Donnelly et al., 1999) then are based partly, but centrally, on the concept that: (1) Serial self-terminating models predict faster increasing positive (target-present) RT-variance functions

(of load) than do negative (target-absent) RT-variance functions. (2) Limited capacity parallel models predict equally increasing positive and negative RT-variance functions (p. 679).

Returning to the former authors' primary interest of assessing the process architecture (serial vs. parallel), consider the first claim above that serial self-terminating models predict a faster increase on target-present than on target-absent trials. It is true that there exist serial models that make that prediction and in fact the so-called standard serial model does (Schneider & Shiffrin, 1977; Townsend & Ashby, 1983).¹ Nevertheless, it is possible to construct serial models that do not make that prediction, and therefore it would be incorrect to infer that serial self-terminating models must make that prediction.

More critically, the second claim that limited capacity parallel models must predict equally increasing standard deviations is seriously flawed. As Townsend and Ashby (1983, pp. 199–201) state, limited-capacity parallel models are capable of predicting a wide variety of variance relationships across stopping rules. In fact, any limited-capacity parallel model that is equivalent to a serial model that makes the present unequal variance-slope predictions must predict exactly the same variance characteristics.

For instance, consider the serial model with constant exponential processing rate u and some distribution on processing paths through the items. Then, this standard type of serial model assumes that each match time is distributed with theoretical frequency function (i.e., probability density) $Ue^{-U|t|}$, with individual mean times of $1/U$. The exhaustive mean RT prediction is then simply n/U , and the self-terminating prediction is $(n + 1)/(2U)$. This model further makes the unequal variance-slope predictions for exhaustive versus self-terminating trials cited by Donnelly et al. (1999; see, Schneider & Shiffrin, 1977; Townsend & Ashby, 1983, pp. 192–201).

However, the classic fixed-capacity parallel model with reallocation mentioned earlier (e.g., Townsend, 1971; Townsend & Ashby, 1983, pp. 88–89) with rates that capture the path distribution is mathematically equivalent to this serial model and therefore also makes the identical variance prediction. These facts suggest that variance too does not necessarily decide between architectures, without ancillary assumptions, any more than do the means.

We next illustrate the variety of variance predictions possible in parallel processing. A functional equation is offered that, if satisfied, does predict the relation between positive and negative trial standard deviations posited by Donnelly et al. (1999). Then, issues regarding generic predictions of parallel models are considered.

One interesting class of parallel models that do *not* conform to the authors' predictions has the property that the variance for negative trials increases considerably faster than that for positive trials. Thus, consider a parallel model with stochastically independent, exponential processing on the items, with rate $V(n)$ on all items, where $V(n)$ may be a function of load, n . This model is interesting because an intriguing qualitative contrast of self-

terminating versus exhaustive processing variances does not depend on whether the system is limited, unlimited, or super capacity—that is, whether $V(n)$ decreases, is constant, or increases as a function of load, n . The self-terminating variance is shown in Equation 1 and the exhaustive variance in Equation 2.

$$\text{VAR}[\text{RT} | \text{S} - \text{T}] = 1/[V(n)]^2 \quad (1)$$

and

$$\text{VAR}[\text{RT} | \text{EXH}] = \{1/[V(n)]^2\} \sum_{j=1}^n (1/j)^2. \quad (2)$$

It may be noted that we purposely dispensed with the residual time (i.e., durations stemming from outside the search process of interest) random variable here because it plays no role of interest.

The self-terminating variance prediction in Equation 1 derives from the fact that, in parallel processing, the probability distribution for the self-terminating time of the target comparison is just the marginal individual distribution for that particular item. Since all the items in this model are processed independently, with identical distributions, each of which is exponential with mean $1/V(n)$ and variance $1/[V(n)]^2$, the prediction is immediate.

The exhaustive variance in Equation 2 is found, because of independence, to be the sum of the variances of each successive completion time random variable. At the j th stage (i.e., when $j - 1$ items have been completed), there are $n - j + 1$ items still being processed, each with mean rate $V(n)$ and the successive intercompletion times are independent (Townsend & Ashby, 1983). The intercompletion time random variable at stage j is exponentially distributed with a rate $(n - j + 1)V(n)$ (i.e., the sum of the individual rates due to parallel processing at stage j), mean $1/[(n - j + 1)V(n)]$ and variance $1/[(n - j + 1)V(n)]^2$. Then, the overall variance is just the sum over j of the latter individual variances. The upshot is that the ratio of the exhaustive-to-self-terminating variances (i.e., Equation 2 divided by Equation 1) increases as a function of n , whatever the capacity assumption—that is, whatever the behavior of $V(n)$ as n varies!

Of course, if processing is exhaustive even on positive trials, and the distributions on processing times are the same on negative and positive trials, the rate of change of the variance as a function of load will be the same as claimed by the authors. But, of course, that perpetrates an unfair comparison, because it is being assumed, in the serial case, that processing is self-terminating.

What mathematical conditions must be met by parallel models in order to satisfy the claim of Donnelly et al. (1999), even when processing is self-terminating on positive trials? This question can be answered by constructing a functional equation that captures the necessary and sufficient conditions for parallel, self-terminating, and exhaustive variance functions of display size to possess equal slopes. For simplicity, assume that processing is parallel and stochastically independent on the separate items (the latter assumption is easily dispensed with but

at the expense of more complexity). Assume also that, although the negative processing rates (i.e., comparison of the probe to distractors) are the same in both positive and negative trials, positive matches can have different rates (an often made and rather reasonable assumption, e.g., see, Bamber, 1969; Krueger, 1978; Snodgrass & Townsend, 1980).

Let T_i stand for the random variable for a negative match in position i and Z_j stand for a positive match in position j . Then,

$$\begin{aligned} \text{VAR}[\text{RT} | \text{EXH}] &= \text{VAR}[\text{MAX}(T_1, T_2, \dots, T_n)] \\ &\equiv [1/(n)] \sum_{j=1}^n V[Z_j] = \text{VAR}[\text{RT} | \text{S} - \text{T}] \\ &= \text{AVE. VAR}[\text{SINGLE TARGET}], \quad (3) \end{aligned}$$

gives the required functional relationship that renders the positive and negative variance slope equal as posited by Donnelly et al. (1999).² The three-lined equal sign in Equation 3 simply indicates that the left side for exhaustive processing has to be equivalent to the right side for self-terminating processing. Observe that the variance for a positive trial is the average variance of an individual positive match, since the different positions can be associated with different processing rates. It is possible to write a general integral formula for the exhaustive side in Equation 3 (see Townsend, 1990a), but it would be of little assistance without further assumptions. However, an intuitive version is found with exponential distributions as we now see.

Suppose all of the negative matches, for a given load, n , have rate $V(n)$, whereas any positive match takes place with rate $W(n)$. Then the functional equation established in Equation 3 becomes Equation 4

$$\begin{aligned} \text{VAR}[\text{MAX}(T_1, T_2, \dots, T_n)] &= [1/V(n)]^2 \sum_{j=1}^n \frac{1}{j^2} \\ &= 1/W(n)^2 = \text{VAR}[\text{Single Target Match}]. \quad (4) \end{aligned}$$

As a special case of Equation 4 of some interest, let $V(n) = V$, a constant, meaning that the model is unlimited capacity with regard to negative matches (see, e.g., Townsend & Ashby, 1983, pp. 83-84). Because this model predicts a relatively rapid increase in variance (though not so fast as for a standard serial model), it would not be possible to let the rate for the positive match be constant, for that would predict a constant variance on positive trials! In fact, it is straightforward to see from Equation 4 that $W(n)$, the positive match rate, must decrease fairly dramatically—that is, be limited in capacity, as n grows

$$W(n) = \left(\frac{1}{V^2} \sum_{j=1}^n \frac{1}{j^2} \right)^{\frac{1}{2}},$$

which means that the positive match rate decreases significantly as n increases. It is typically found that positive matches are faster than negative matches (e.g., Bamber,

1969; Krueger, 1978), but, in principle, positive matches could be faster than negative but also decrease faster as a function of load.

In any event, it can be seen from analyses of Equations 3 and 4 that it would be only very special kinds of parallel models that allow self-termination on positive trials that would predict equal sloped variance functions.

Before leaving the issue of architecture and variances, it can well be asked whether there are any invariant characteristics of variances associated with parallel processing. As with means, it appears that, if capacity relations are not restricted, then virtually anything goes. However, what about what has been called (in analogy to the standard serial model, see Note 1) the *standard parallel type of model*, in which processing is assumed to be stochastically independent on the various items and capacity is unlimited? This is the kind of parallel model originally tested by Sternberg (1966; Townsend, 1972; Townsend & Ashby, 1983). Unlimited capacity implies that the individual item distributions are constant over load, n , although they might vary across location or item per se. The variance statistic from self-terminating trials will then be constant across load in any such model, itself a readily falsifiable prediction. However, exhaustive variance behavior is even more interesting.

In general, the processing time variance from exhaustive trials can increase or decrease, and, if the distribution is symmetric, it appears to always decrease from $n = 1$ to $n = 2$ (Townsend, 1990a). The latter study speculates on the basis of the previous theorem as well as on a number of examples that exhaustive parallel processing is inclined to deliver decreases in variance as n increases. Furthermore, symmetric and negatively skewed distributions seem to offer evidence that variance decreases across n . In the examples so far, variance has increased with n only when the distribution was very positively skewed (e.g., the exponential distribution).

In addition, a new more powerful mathematical approach regarding variability based on quantile spread statistics (Townsend & Colonius, 2001) finds results that appear broadly compatible with the early results on variances. That is, exhaustive processing tends to produce decreases in variability.³ This tendency is specified quantitatively. Probability frequency functions (technically, *probability density functions*) that are always increasing and, interestingly, symmetric U-shaped distributions, must predict decreases in variability as n increases. It appears that negatively skewed distributions will also largely evidence decreases in variability with increases in n .⁴ However, extremely positively skewed distributions, such as the exponential distribution can predict increases in quantal range as n increases. While, in principle, unimodal symmetric frequency functions (i.e., like the normal distribution) can predict increases, the individual ones we have probed so far have yielded decreases. A very intriguing specific distribution, the *double exponential distribution*, with cumulative probability function given by $F(x) = \text{Exp}[-\text{Exp}(-x)]$, turns out to possess just enough

positive skew to exactly balance the tendency to yield decreases, which therefore predicts no change of variability with n at all!

Most psychologists who have been informally interviewed report that they would be surprised to find data with decreasing RT variance as load n increases. However, if capacity is unlimited and the channels are independent (standard parallel processing), this may not be uncommon behavior. Nevertheless, quantitative investigation of popular distributions used in RT modeling (e.g., Poisson counting models, Luce, 1986; Townsend & Ashby, 1983; or ex-Gaussian, Ratcliff, 1978; or those produced by random walk or diffusion models, Link & Heath, 1975; Ratcliff, 1978; Smith, 1990) is called for before firm conclusions can be drawn in this regard. In any event, there may be hope on the horizon for at least separating, through powerful qualitative tests, certain large classes of standard parallel models from standard kinds of serial models and perhaps from less orthodox (e.g., limited capacity and/or dependent) parallel models as well.

Varying n and Assessing the Stopping Rule

A final matter regards asymmetries in the ways in which various dependent variables impact different theoretical properties. Thus, as indicated earlier, mean RT slopes (especially slopes of any magnitude greater than 0) are known to be ineffectual with regard to the parallel-serial distinction, and we have demonstrated RT variances to also be lacking in this regard. In fascinating contrast, RT slopes and position effects (RT differences according to the location of the target item) have been mathematically demonstrated to be a powerful tool in testing for self-terminating versus exhaustive processing. The latter point stresses the value of quantitative meta-theoretical investigations in not only bringing into question problematic criteria concerning psychological mechanisms but also their value in substantiating valid criteria.

It required a patient line of theoretical and methodological research to first reveal parallel-serial testing dilemmas in popular experimental testing paradigms that later helped lead to more powerful means of performing experimental tests. A similar kind of effort showed that the same statistic (i.e., slope of mean RTs as a function of load, n) that was problematic with regard to the parallel-serial question, actually provided a strong test in the case of self-terminating versus exhaustive processing (Townsend & Ashby, 1983; Townsend & Colonius, 1997; Townsend & Roos, 1973; Townsend & Van Zandt, 1990; Van Zandt & Townsend, 1993). Those studies also reveal that target position effects provide powerful criteria for self-terminating versus exhaustive processing. Whether slope of variances and variance as a function of target position will be equally helpful with respect to the stopping rule question is as yet undetermined.

Conclusion

The specific lesson here must be that both the mean and variance are functions not only of architecture but of equal importance, of capacity and dependence relations, and

that great care must be taken in conclusions regarding architecture from the data they provide. Nevertheless, they can and do reduce the class of possible explanatory models when taken in concert; the story is just not always quite simple. And, when insufficient care is taken with regard to the logical structure of model testing relationships, it will be easy to miss the distinction between valid and invalid test methodologies. The latter is, as noted above, likely to be found in many areas of cognitive and perceptual psychology that have not been subjected to the intensive meta-theoretical analyses that have been visited on parallel-serial methodology. This fact renders the present concerns of general import.

REFERENCES

- ASHBY, F. G., TEIN, J.-Y., & BALAKRISHNAN, J. D. (1993). Response time distributions in memory scanning. *Journal of Mathematical Psychology*, *37*, 526-555.
- 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.
- BAMBER, D. (1969). Reaction times and error rates for "same"–"different" judgments of multidimensional stimuli. *Perception & Psychophysics*, *6*, 169-174.
- DONNELLY, N., FOUND, A., & MÜLLER, H. J. (1999). Searching for impossible objects: Processing form and attributes in early vision. *Perception & Psychophysics*, *61*, 675-690.
- EGETH, H. E. (1966). Parallel versus 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.
- ENNS, J. T., & RENSINK, R. A. (1990). Influence of scene-based properties on visual search. *Science*, *247*, 721-723.
- KENDALL, M. G., & STUART, S. (1969). *The advanced theory of statistics* (Vol. 1). New York: Hafner.
- KRUEGER, L. E. (1978). A theory of perceptual matching. *Psychological Review*, *85*, 278-304.
- LINK, S. W., & HEATH, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*, 77-105.
- LUCE, R. D. (1986). *Response times*. New York: Oxford University Press.
- MURDOCK, B. B., JR. (1971). A parallel-processing model for scanning. *Perception & Psychophysics*, *10*, 289-291.
- NEUFELD, R. W. J., & McCARTY, T. S. (1994). A formal analysis of stressor and stress-proneness effect on simple information processing. *British Journal of Mathematical & Statistical Psychology*, *47*, 193-226.
- RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1-66.
- 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., & TOWNSEND, J. T. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, *33*, 309-327.
- SMITH, P. L. (1990). A note on the distribution of response times for a random walk with Gaussian increments. *Journal of Mathematical Psychology*, *34*, 445-459.
- SNODGRASS, J. G., & TOWNSEND, J. T. (1980). Comparing parallel and serial: Theory and implementation. *Journal of Experimental Psychology: Human Perception & Performance*, *6*, 330-354.
- STERNBERG, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652-654.
- STERNBERG, S. (1998). Discovering mental processing stages: The method

- of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science, Vol. 4: Methods, models and conceptual issues* (pp. 739-811). Cambridge, MA: MIT Press.
- TOWNSEND, J. T. (1969). Mock parallel and serial models and experimental detection of these. In *Proceedings of the Purdue Centennial Year Symposium on Information Processing* (pp. 617-628). West Lafayette, IN: Purdue University, School of Electrical Engineering.
- 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 & Statistical Psychology*, **25**, 168-199.
- 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. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology*, **28**, 363-400.
- TOWNSEND, J. T. (1990a). *A potpourri of ingredients for horse (race) soup*. (Tech. Rep. No. 32). Bloomington, IN: Indiana University, Cognitive Science Program.
- TOWNSEND, J. T. (1990b). 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., & ASHBY, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.
- TOWNSEND, J. T., & COLONIUS, H. (1997). Parallel processing response times and experimental determination of the stopping rule. *Journal of Mathematical Psychology*, **41**, 392-397.
- TOWNSEND, J. T., & COLONIUS, H. (2001). *Variability of the MAX and MIN statistic: A theory of the quantile spread as a function of sample size*. Manuscript submitted for publication.
- TOWNSEND, J. T., & DIETRICH, A. (1996). The relationship of variance to interaction contrast in parallel systems factorial technology. *British Journal of Mathematical & Statistical Psychology*, **49**, 211-223.
- TOWNSEND, J. T., & ROOS, R. N. (1973). Search reaction time for single targets in multiletter stimuli with brief visual displays. *Memory & Cognition*, **1**, 319-332.
- 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* (pp. 469-489). Amsterdam: North-Holland.
- VAN ZANDT, T., & TOWNSEND, J. T. (1993). Self-terminating versus exhaustive processes in rapid visual and memory search: An evaluative review. *Perception & Psychophysics*, **53**, 563-580.
- VORBERG, D. (1977, August). *On the equivalence of parallel and serial models of information processing*. Paper presented at the Tenth Annual Mathematical Psychology Meetings, University of California, Los Angeles.
- WOLFE, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, **9**, 33-39.

NOTES

1. It is difficult to make a theoretical point without technical detail, but since a note of this sort must remain brief, let us just take the standard serial model as assuming identical, independent probability distributions on the time for each item to be searched. (An exponential version of the standard serial model is outlined later in the paper.) This definition can be generalized but this is sufficient for the present purposes. (See Townsend & Ashby, 1983, for further detail.)
2. In fact, we make the slightly stronger assumption that the self-terminating and exhaustive variances are actually equal.
3. We also developed the theory for minimum response time, which, in parallel processing, is associated with so-called race models. However, the emphasis in the present paper is on exhaustive and ordinary self-terminating processing.
4. We are here using the traditional terms of skewness and kurtosis (just below) to indicate asymmetry and "sharpness" of the frequency functions. However, our approach employs more powerful definitions of these concepts than is the case with the somewhat problematic earlier definitions by Pearson and others (see, e.g., Kendall & Stuart, 1969).

(Manuscript received March 1, 2000;
revision accepted for publication November 22, 2000.)