ISB2 43103

AU1 a0005

Information Processing Architectures: Fundamental Issues

James T Townsend and Haiyuan Yang, Indiana University, Bloomington, IN, USA
Trisha Van Zandt, Ohio State University, Columbus, OH, USA

Abstract

abspara0010 This chapter starts with a brief history of information processing architectures, emphasizing a traditional experimental paradigm, the additive factors method, and the classic problem of model mimicry. Several solid approaches to identifying mental architecture are introduced with a discussion of the necessary assumptions for these tests. Other fundamental issues attached to information processing systems are also presented. A brief discussion with regard to challenges for the future concludes the chapter.

p0010 In the late nineteenth century, budding experimental psychology began to investigate perceptual and mental operations that would be considered part of the contemporary purview of cognition. Such interests reappeared in the 1960s in the emerging information processing approach to elementary cognition. Over the last several decades of the twentieth century, considerable progress has been made in this tradition through the combination of mathematical modeling and experimental design, an approach becoming known as metamodeling. This section reviews a number of selected findings within this approach, which is relatively novel in the social sciences. One benefit of the strategy is that potential dilemmas involving model mimicking, where models based on strikingly different principles can imitate each other's predictions, are brought to light. In particular, problems and experimental solutions associated with testing among parallel (simultaneous) versus serial (one-at-a-time) processing, stopping rules (logical basis for cessation of processing), limited versus unlimited versus Supercapacity (effects of increasing workload on processing speed), and dependence (issue of stochastic independence of channels or events) relationships are described. The foregoing issues are fundamental in almost all information processing situations since virtually any operating system must take them into account. The section concludes by pointing to future challenges within this approach.

p0015 The term 'information processing architecture' refers broadly to the arrangement of mental subsystems that are hypothesized to be active in the performance of one or more psychological tasks. For instance, the simplest, most prototypical, and opposed types of architectures are serial (one-at-a-time) versus parallel (simultaneous) arrangement of two or more separate subsystems or processes. More complex arrangements are mentioned below (*see Additive Factor Models*). Further, there are a number of other aspects of perceptual and cognitive processing that are often included under the 'architecture' label including the questions of the basis on which cognitive processes will cease, various kinds of independence and dependence, and processing capacity. These will be described and discussed.

p0020 An early researcher especially pertinent to this review was Donders, a nineteenth-century Dutch physiologist. Donders believed that he could uncover the durations taken by various thought processes through his method of subtraction. The method of subtraction was based on the idea that complicated mental activities were compounded in a simple sequential fashion from less complex parts. Let mean response time be written as $\overline{RT}$ and response time in general as simply RT. Then, for instance, the scientist might engage the subject in a task requiring both perception and decision and compare $\overline{RT}$ from that condition with $\overline{RT}$ from a task requiring only perception. The difference in $\overline{RT}$s was interpreted as the mean duration of the decision operation.

p0025 The issues selected for review here seem elemental in the following sense. The construction of almost any system intended to carry out the processing (e.g., detection, search, comparison, recognition, recall, analysis of various kinds, and so on) of a finite number of tasks or items would have to make decisions on each of the studied issues. In addition, they are somewhat unique in having been subjected to quite general theoretical, quantitative, and methodological analysis, perhaps more than any other such concepts in the field. The issues will be introduced in the context of a popular experimental paradigm and then discussed in more detail.

An Experimental Paradigm and the Major Issues s0010

p0030 The concepts to be defined below have figured prominently in studies on short-term memory and visual display search and we shall employ that type of paradigm for illustration. Considerable impetus was given to the information processing approach by several pioneering studies in the 1960s focusing on short-term memory and visual processing, using RT as the dependent variable ([bib2]Atkinson et al., 1969; [bib10]Egeth, 1966; [bib28]Sternberg, 1966). In the roughly 45 years of the interim, scores and perhaps hundreds of studies in memory and visual search have been carried out. We focus on [bib28]Sternberg's (1966) short-term memory search paradigm for illustration. Short-term memory is specified operationally by the tasks requiring the retention of a small number of items for anywhere from a few seconds to several minutes. In this paradigm, a varying number (less than or equal to the number that can be maintained in short-term memory without error) of items, for instance, randomly arranged letters, is presented to the experimental participant. This is called the memory set. Then, a few seconds later, that participant is presented a so-called probe item and the task is to indicate very quickly, but at a very high level of accuracy, whether or not the probe was contained within the memory set. As introduced above, RT is the primary dependent variable of interest here, although a substantial history of work with accuracy also exists.

p0035 The first, and in some ways the most critical question, referred to the temporal structure of comparison of the probe item with the memory items. A hallmark of many early information processing models was seriality of operation, that is, each subsystem or component could operate only when preceding mechanisms were completed and only one could operate at any given time. Thus, *seriality* has the connotation that no overlap of processing times occurs, in addition to the sequentiality of starting times. This latter type of constraint is referred to as *discrete flow*. The alternative possibility arose early on that all the items might be searched simultaneously, that is, in parallel. This issue of processing architecture can refer to the manner of treatment with regard to items within a single stage of processing, for instance memory search, as in the present case, or to the arrangement of larger components of the system, such as stimulus encoding versus memory storage.

p0040 Of course, arrangements other than serial versus parallel are possible, for instance, a system being serial part of the time and parallel part of the time. When all operations under focus begin at the same time but can finish at different times, the system (or mode of processing) is of course simply parallel. However, when operations can feed into other components, but with overlapping processing times, the system is said to have continuous flow, rather than discrete flow. A great deal has been learned about discrete-flow architectures of considerable complexity over the past two decades (*see Additive Factor Models*). We will confine ourselves to serial and parallel systems and other processing characteristics that these can involve.

p0045 The primary RT durations under study in the memory search task are those associated with comparison of the probe with the memory items. Nonetheless, other time intervals, such as those associated with early sensory processing and late motor components, must also be taken into consideration. It is usual, but by no means uncontested assumption, that these so-called residual times are serially arranged with the process under study and are also stochastically independent of the latter (Dzhafarov, 1992; Luce, 1986; Townsend and Ashby, 1983).

p0050 Besides the architecture issue, independence versus dependence among the processing times of the items is also an important question. For instance, in serial processing, if an earlier item takes more time than usual, a subsequent item might consume less duration due to more preparation time, resulting in a negative dependence among processing times. On the other hand, positive dependences are possible. An example would be a positive dependence created somewhat artificially if *attention* available on a trial were a random AU2 variable for then items would all tend to be processed fast or slow more or less in unison. Analogously, parallel systems can either be constituted with independent channels (e.g., Bundesen and Habekost, 2008; Eidels et al., 2010) or with positively (e.g., Townsend and Wenger, 2004b; Wenger and Townsend, 2001) or negatively (e.g., Usher and McClelland, 2003) interactive channels.

p0055 Another critical concept is that of *workload capacity* (or 'capacity' for short if the meaning is clear), which refers to how processing times are affected by the number of things to be worked on (Townsend and Ashby, 1983). This is most easily illustrated with a version of parallel processing where the system slows down when the number of items that are being processed increases (i.e., limited capacity). However, limitations in capacity that are indirect, even with serial processing, can be conceived. For instance, a serial processor might speed up as it goes through the items, due to warm-up effects, or slow down due to inertial factors or fatiguing of the processor. Even though capacity and independence are logically separate notions, they can interact. For instance, an important type of parallel system, one that can mimic serial processing, assumes that as the processing of each item is completed, its processing capacity is reallocated to remaining items (Townsend and Ashby, 1983). This obviously affects the overall RT, but also creates a positive dependence among the item processing times. Townsend and Wenger (2004b) build a theory and associated methodology featuring the capacity theme.

p0060 Yet another important notion is that of stopping rule (e.g., Sternberg, 1966; Townsend, 1974). Depending on the task, it may or may not be necessary for the participant to process all of the items in order to make a correct response. In the memory search paradigm, if the probe item is present and located in the current stimulus set, the processing can cease at that instant, without finishing the remaining items. This possibility is known as self-termination. However, since short-term memory search consumes only a few hundred milliseconds it is possible that the system nevertheless completes all items. This event is called *exhaustive* processing. On probe-absent trials, it is necessary to process all of the memory items in order to be sure of correctly making a 'no' response, that is, exhaustive processing must occur. In some experimental designs, all the items are probes. These need not be physically identical. For instance, the task might be to say 'yes' if any of the items is a vowel and a target-present trial could contain all vowels. This latter case permits the possibility of *first-terminating* or *minimum time* processing. Of course, it is an empirical question as to whether any kind of self-termination can actually take place in high-speed perceptual or cognitive operations, and must be addressed experimentally in each case.

p0065 Space constrains the scope of mathematical detail in the present article, but we provide some fundamental, if simplified, depiction. Although the serial and parallel classes of models both contain an infinite class of possibilities, the serial notion has traditionally been attached to a particular serial model that assumes identical processing time random variables on each item, independent of the number of items in memory (i.e., the load) and of the order of processing the items. It is also often assumed that the individual processing times are themselves mutually stochastically independent. We call this the *standard serial model*. Let the density of processing time for each item be designated $f(t)$ and that for the independent residual processing time be $g(r)$, where T and R are the respective random variables for t and r. Then the density on a probe-absent trial, $p(\mathrm{RT}=t)$, or $p(t)$ for short, is just the mathematical convolution of the n processing densities and

AU3 the residual time density: $p(t|\mathrm{EX}) = f_1 * f_2 * \ldots * f_n$, where each f_i, $i = 1, \ldots, n$, is a replica of. The expected or average theoretical RT is $E(\mathrm{RT}) = n \cdot E(T) + E(R)$. Notice that processing is exhaustive, as designated by EX in the left-hand side, in this case. In the case of self-terminating (ST) processing on a probe-present trial, the number of items finished before the probe is found is itself a random variable. Under usual conditions, the probability that the probe will be found in the *i*-th processing position is just $1/n$, so the average density for this case is just $E[p(T|ST)] = (1/n) \sum_{j-1}^{n} f_1 * f_2 * \ldots * f_{i-1} * f_i$ times. Similarly, the expected RT, with $\overline{\mathrm{RT}}$ as the sample statistic, is just $E[(\mathrm{RT}|\mathrm{ST})] = (1/n) \cdot \sum_{j-1}^{n} E(T) + E(R) = [(n+1)/2] \cdot E(T) + E(R)$, which yields the time-honored result that this line has half the slope of that for the exhaustive serial RT.

p0070 The concept of parallel processing, like serial processing, is usually confined in the experimental literature to a very special kind of parallel processing. Processing is often said to be parallel if average RT is invariant across increases in load. When exhaustive processing has to occur, as in the probe-absent trials of memory search, this stipulation actually implies a very unusual kind of parallel processing, unlike the situation with the standard serial model. For instance, consider a set of parallel models with independent processing times, with the additional provision of unlimited capacity, in the sense that the probability distribution on completion time for each item does not change as *n* is increased. This class of models serves as a prototype of parallel processing in an analogous way to the standard serial model class just presented. Let us call this class the standard parallel model.

p0075 Let $g(t)$ be the fixed density function on processing time for each of the n items. Then $E[\mathrm{RT}|\mathrm{EX}] = \mathrm{E}\{\mathrm{MAX}[T_1, T_2, \ldots T_{n-1}, T_n]\} + E(R)$, where R is as above and T_i is the processing time of the item in position i. We drop reference to R which plays no role in the function shape, letting PT stand for processing time, that is, $\mathrm{PT} = \mathrm{RT} - R$. Then $E[(\mathrm{PT}|\mathrm{EX})] = \int_{t'-0}^{\infty} [1 - G(t')]^n \mathrm{d}t'$, where $G(t)$ is the cumulative distribution function associated with g. It is straightforward to show from this formula that mean RT is indeed an increasing, concave function of n (Townsend and Ashby, 1983, p. 92). Note that $g(t)$ is invariant across values of n. Hence, many investigators use a very restrictive and typically unrealistic criterion for parallelism when they demand flat rather than increasing functions $\overline{\mathrm{RT}}$ of n, even under exhaustive processing conditions.

p0080 It can also be observed that if mean exhaustive $\overline{\mathrm{RT}}$s are flat functions of n, then mean ST (again, a single position exists in the memory list that contains the probe) times would actually be decreasing, within the same model – a strong prediction which apparently has never been checked in studies using this logic. Thus, in the rare cases where flat exhaustive $\overline{\mathrm{RT}}$ functions are actually found with exhaustive processing, the implications are quite strong and in favor of very supercapacity parallel processing. Models that can make such predictions are considered in Townsend and Ashby (1983, Chapter 4).

p0085 In contrast, in the case of ordinary ST processing when a single position contains the probe, standard parallel models do predict a flat $\overline{\mathrm{RT}}$ function. This can be readily intuited, since only the channel processes the probe matters in ST processing, and it is independent of all the rest (and is of unlimited capacity into the bargain). Finally in this model, mean first-terminating times should decrease with n.

Important Related Research Topics s0015

p0090 Although the issues considered here are applicable to virtually any information processing system or task, there are some classical or contemporary subject areas for which they are particularly germane.

p0095 When applied within the confines of a single subprocess, such as search within a memory or display list, the concept of attention is evoked. How attention is deployed, when it ceases, whether its application is independent across the various items, and its strength as a function of the workload, all call upon the critical processing issues introduced above.

p0100 The topic of *automatic processing* traditionally discussed within the realm of attention, likewise overlaps our present issues (e.g., Schneider and Shiffrin, 1977). The most obvious correlate within our processing issues is that of capacity. Although writers are sometimes rather vague concerning detailed quantitative accounts of what automaticity means with regard to workload capacity, it seems apparent that parallel processing is definitely mandated. In addition, we have proposed that capacity should at least be at the unlimited level. That is, each subsidiary task should see its operations proceed at the same speed as if it were functioning alone. These and related matters are discussed further in Neufeld et al. (2007).

p0105 Another traditional theme, but also seeing resurgence, is that of perceptual and cognitive *holism*, a topic of long standing in philosophy and in the twentieth century, of *Gestalt psychology*. Although much has been learned about psychological holism, there has been little accomplished with regard to definitions, theory, and experimental investigation of information processing characterizations of dynamic holistic operations. In that spirit, a set of working axioms which portray holistic perception in terms of the present issues have been proposed and developed (e.g., Wenger and Townsend, 2001; Wenger and Ingvalson, 2003; Fific and Townsend, 2010). The accompanying essentials of holistic perception can be informally expressed as involving highly interactive, supercapacity, exhaustive, parallel processing.

The Mimicking Dilemma of Serial versus Parallel Processing s0020

p0110 As mentioned earlier, one of the driving forces behind mathematical metamodeling in this area was the discovery that mathematical representations of diametrically opposed psychological principles could nevertheless sometimes mimic one another, even to the point of mathematical equivalence of their defining probability laws (Murdock, 1971; Townsend, 1971; Townsend and Ashby, 1983). A historical account of this dilemma and its resolution is offered by Townsend et al. (2011). Hence, we will take some time here to outline the state of the art with regard to such questions within the parallel versus serial question.

p0115 Consider a model for the processing times of *n* items under fixed experimental conditions. With regard to the parallel–serial issue, suppose no assumptions are made other than the probability mixture of generalized convolutions in the case of seriality and joint distributions on processing times in the case of parallelism. Then the parallel and serial classes of models are equivalent, in the sense that mappings can be provided that homeomorphically (this is mathematical jargon for a one-to-one, and continuous to-and-fro function relating the two types of models) carry the serial probability distribution into the parallel probability distribution and vice versa (Townsend and Ashby, 1983).

p0120 Nevertheless, over time, accumulating theoretical results have demonstrated that if the scientist is willing to make further restrictive, but sometimes still very general (and reasonable), assumptions about the models, and/or more complex experimental designs are utilized, the parallel–serial issue can be decided. For instance, certain rather fundamental differences between serial and parallel processing can be explored in experimental methods designed to exploit those differences (Townsend and Ashby, 1983; Townsend and Wenger, 2004a).

s0025 Factorial Methodologies for Architectures and Systems Factorial Technology

p0125 One of the most promising and general approaches to identifying mental architecture (i.e., serial versus parallel processing) is that based on the notion of *selective influence* of experimental factors, a notion first employed in tests of strict seriality by Sternberg (1969) in his well-known additive factors method (*see Additive Factors Models*). All factorial methodologies, like the original Sternberg strategy, depend on the selective influence assumption, namely, that distinct experimental factors affect distinct processing components (i.e., subsystems), the assumption of selective influence. It can be assumed that $\mathrm{RT}(X+\Delta X,\ Y)$ refers to the case where the X factor has prolonged RT but Y is at base level, and so on. Basically, the fundamental statistic for the original method and for most extensions was the *mean interaction contrast* (MIC). The MIC is defined as

$$\begin{aligned}\mathrm{MIC} &= \overline{\mathrm{RT}}(X+\Delta X, Y+\Delta Y) - \overline{\mathrm{RT}}(X+\Delta X, Y)\\ &\quad - [\overline{\mathrm{RT}}(X, Y+\Delta Y) - \overline{RT}(X, Y)]\end{aligned}$$

p0130 Schweickert (1978) in his *latent mental network theory* contributed the first major extension of the additive factors method involving more complex architectures under the assumption of selective influence. Townsend and Ashby (1983) found that the MIC distinguished parallel and serial stochastic models when selective influence was assumed and Schweickert and Townsend (1989) produced general theorems for Schweickert's latent networks, within a stochastic setting, assuming exhaustive processing (see *Additive Factor Models*; AU4 43069).

p0135 Although the early theorems in all this work were accomplished in the context of exhaustive processing, analogous results can be found in the case of ST and first-terminating processing times (e.g., Townsend and Nozawa, 1995; Townsend and Wenger, 2004a). Because Sternberg's original ideas inherent in his additive factors method have been extended in so many new directions, it has been suggested that the general approach be referred to as systems factorial technology (Townsend and Thomas, 1994). For instance, one novel strategy has been to enlist entire RT distributions in providing more powerful tests of parallel versus serial processing or other related architectures (Dzhafarov and Schweickert, 1995; Roberts and Sternberg, 1993; Townsend, 1990; Townsend and Ashby, 1983; see also Balakrishnan, 1994). For instance, in analogy to the MIC, the scientist can form a contrast function composed of the double difference (corresponding to the double difference in mean RTs in the usual case) of cumulative distribution functions. This new statistical function turns out to be very helpful in assaying mental architecture (Townsend and Nozawa, 1995). Another example of the use of entire distributions will be reviewed below.

The assumption of selective influence is critical to the legitimacy of systems factorial technology. Much has been learned in recent years about its foundational underpinnings, what may go awry if it is violated, and about certain experimental indications of its failure (Dzhafarov, 1997; Dzhafarov and Kujala, 2010; Townsend and Schweickert, 1989; Townsend and Thomas, 1994). p0140

Discriminability Results on Stopping Rule, Capacity, and Dependence s0030

The question may be raised as to whether processing issues other than parallel versus serial processing also suffer significant problems in identification within the search or related paradigms. With regard to the stopping rule, mathematical investigations have shown, somewhat ironically, that the same types of memory search data that were incapable of deciding the parallel–serial question could in many instances prove that processing was ST rather than exhaustive. Interestingly, it is more difficult to prove that processing is exhaustive in that ST models can mimic exhaustive processing but not vice versa (Townsend and Colonius, 1997; Van Zandt and Townsend, 1993). p0145

Moreover, recent theoretical and empirical discoveries utilizing the entire RT distributions rather than means provide much strengthened tests of architecture (parallel versus serial) and in addition allow one to firmly distinguish stopping rules at the same time. In fact, within systems factorial technology and assuming selective influence at the distributional level, Townsend and Nozawa 1995 (see also Townsend and Wenger, 2004a) proposed an experimental design that they called the *double factorial paradigm* in which investigators can test architecture and stopping rule as well as capacity within the same block of experimental trials. p0150

Additionally, the important matter of workload capacity seems to possess little in the way of mimicking predicaments. Whether capacity increases, decreases, or is invariant, in the face of workload alterations is immediately captured by the current statistical measures. p0155

The most difficult issue of the ones under discussion is that of stochastic dependence, even though it plays a vital role in processing systems. Interestingly, at this point in time, dependence is more readily and more directly assessed within experimental p0160 designs based on accuracy rather than RT (Ashby and Townsend, 1986; Kadlec and Townsend, 1992).

s0035 Challenges for the Future

p0165 The bringing together of mathematics and experimental methodology has lead to a useful stratagem that we call metamodeling. Systems factorial technology provides an example of metamodeling. Metamodeling facilitates the development of experimental methodologies that are capable of testing broad classes of models, rather than highly specific models, against one another. This approach is especially helpful when attempting to settle well-defined issues, for example, diametrically opposed concepts such as parallel versus serial processing. It may be more difficult to use this strategy with highly complex and detailed models of phenomena (e.g., Van Zandt and Ratcliff, 1995).

p0170 Metamodeling has arguably led to striking advances in the ability to decide experimentally a number of elementary, but major, issues in the purview of human information processing over the last several decades of the twentieth century. Some of these were outlined above. However, many challenges remain. For instance, long-standing challenges relate to the somewhat pesky residual time component introduced earlier. It is difficult to know in many cases how successful our various strategies can be until its nature is firmly tied down. As noted earlier the residual time is usually assumed to be in series with, and stochastically independent of, the other processing components (Luce, 1986; Smith, 1995; Townsend and Ashby, 1983). However, this assumption may be and has been challenged. Thus, it has been contrasted with the extreme opposite assumption of perfect correlation, but still within a series arrangement of the residual component with the other stages of processing (Dzhafarov, 1992; see also a specific case of apparent nonindependence, Diederich and Colonius, 1991).

p0175 The most time-honored approach has been to attempt to separate the residual time component from those under study employing Fourier analysis, assuming stochastic independence of these components. This has proven to be a tricky and arduous strategy. Luce (1986) and Sheu and Ratcliff (1995) provide useful reviews and commentaries on this approach. Another kind of powerful, if often difficult, approach regarding the residual time component has been to prove that experimental predictions hold true regardless of its presence (Dzhafarov and Schweickert, 1995; Roberts and Sternberg, 1993; Townsend and Ashby 1983).

p0180 The residual time problem also abuts a more general possibility, namely that processes in the same forward sequence may not satisfy discrete flow, but rather obey continuous flow properties. For instance, almost all systems based on differential and integral equations would fall into the continuous flow class. Some intriguing progress, both theoretical and empirical, has been made on certain subclasses of such systems (McClelland, 1979; Miller, 1993; Schweickert and Mounts, 1998), but much more remains to be done.

p0185 Another evident opportunity for development is the extension of factorial methods in general, and systems factorial technology in particular, to other dependent variables than response times. The most palpable extension would be to accuracy or accuracy and response times. Another interesting and topical subject matter would be *confidence ratings*. In this regard, it has been pointed out that it constitutes a grave error to simply assume that predictions for response times will also apply to other observable variables (Townsend, 1984). Theorems that are appropriate for specific observables within an information processing milieu must be proven and tested in the experimental crucible.

As a final exemplary vein that is offered is that of implementing the concept of a theoretical and methodological *sieve*. Within such a sieve, the investigator first establishes the major processing characteristics of information processing in a cognitive task, and then proceeds to propose and test more computationally detailed accounts that are in accord with the earlier findings. The methodological sieve is related to, but is more general than, the concept of *strong inference* put forth by Platt (1964) (e.g., Massaro, 1998). p0190

See also: Additive Factor Models; 43039; 43069. ED1

Bibliography

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

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

Balakrishnan, J.D., 1994. Simple additivity of stochastic psychological processes: tests and measures. Psychometrika 59, 217–240.

Bundesen, C., Habekost, T., 2008. Principles of Visual Attention. Oxford University Press, Oxford.

Diederich, A., Colonius, H., 1991. A further test of the superposition model for redundant signals effect in bimodal detection. Perception and Psychophysics 50, 83–86.

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

Dzhafarov, E.N., 1997. Process representations and decompositions of response times. In: Marley, A.A.J. (Ed.), Choice, Decision, and Measurement: Essays in Honor of R. Duncan Luce. Erlbaum Associates, Mahwah, NJ, pp. 255–278.

Dzhafarov, E.N., Kujala, J.V., 2010. The joint distribution criterion and the distance tests for selective probabilistic causality. Frontiers in Quantitative Psychology and Measurement 1, 151.

Dzhafarov, E.N., Schweickert, R., 1995. Decompositions of response times: an almost general theory. Journal of Mathematical Psychology 39, 285–314.

Egeth, H., 1966. Parallel versus serial processes in multidimensional stimulus discrimination. Perception and Psychophysics 1, 245–252.

Eidels, A., Donkin, C., Brown, S.D., Heathcote, A., 2010. Converging measures of workload capacity. Psychonomic Bulletin and Review 17, 763–771.

Fific, M., Townsend, J.T., 2010. Information-processing alternatives to holistic perception: identifying the mechanisms of secondary-level holism within a categorization paradigm. Journal of Experimental Psychology: Learning, Memory, and Cognition 36, 1290–1313.

Kadlec, H., Townsend, J.T., 1992. Signal detection analysis of dimensional interactions. In: Ashby, F.G. (Ed.), Multidimensional Models of Perception and Cognition. Erlbaum, Hillsdale, NJ, pp. 181–228.

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

Massaro, D.W., 1998. Perceiving talking faces: from speech perception to a behavioral principle. MIT Press, Cambridge, MA.

McClelland, J.L., 1979. On the time relations of mental processes: an examination of systems of processes in cascade. Psychological Review 86, 287–330.

Miller, J., 1993. A queue-series model for reaction time, with discrete-stage and continuous-flow models as special cases. Psychological Review 100, 702–715.

Murdock Jr., B.B., 1971. A parallel-processing model for scanning. Perception and Psychophysics 10, 289–291.

Neufeld, R.W.J., Townsend, J.T., Jetté, J., 2007. Quantitative response time technology for measuring cognitive-processing capacity in clinical studies. In: Neufeld, R.W.J. (Ed.), Advances in Clinical Cognitive Science: Formal Modeling and Assessment of Processes and Symptoms. American Psychological Association, Washington, D.C.

Platt, J.R., 1964. Strong inference. Science 146, 3642.

Roberts, S., Sternberg, S., 1993. The meaning of additive reaction time effects: tests of three alternatives. In: Meyer, D.E., Kornblum, S. (Eds.), Attention and Performance XIV. MIT Press, Cambridge, MA, pp. 611–654.

Schneider, W., Shiffrin, R.M., 1977. Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Bulletin 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., Mounts, J., 1998. Additive effects of factors on reaction time and evoked potentials in continuous flow models. In: Roberts, F.S., Dowling, C., Theuns, P. (Eds.), Progress in Mathematical Psychology. Erlbaum Associates, Mahwah, NJ, pp. 311–328.

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

Sheu, C.F., Ratcliff, R., 1995. Fourier deconvolution of reaction time data: a cautionary note. Psychological Bulletin 116, 285–299.

Smith, P.L., 1995. Psychophysically principled models of visual simple reaction time. Psychological Review 102, 567–593.

Sternberg, S., 1966. High-speed scanning in human memory. Science 153, 652–654.

Sternberg, S., 1969. The discovery of processing stages: extensions of Donders' method. In: Koster, W.B. (Ed.), Attention and Performance, vol. 2. North Holland Press, Amsterdam.

Townsend, J.T., 1971. A note on the identifiability of parallel and serial processes. Perception and Psychophysics 10, 161–163.

Townsend, J.T., 1974. Issues and models concerning the processing of a finite number of inputs. In: Kantowitz, B.H. (Ed.), Human Information Processing: Tutorials in Performance and Cognition. Erlbaum Press, Hillsdale, NJ, pp. 133–168.

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

Townsend, J.T., 1990. The 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. (Eds.), 1983. Stochastic Modeling of Elementary Psychological Processes. Cambridge University Press, Cambridge, UK.

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., Nozawa, G., 1995. On the spatio-temporal properties of elementary perception: an investigation of parallel, serial and coactive theories. Journal of Mathematical Psychology 39, 321–360.

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

Townsend, J.T., Thomas, R.D., 1994. Stochastic dependencies in parallel and serial models: effects in systems factorial interactions. Journal of Mathematical Psychology 38, 1–34.

Townsend, J.T., Wenger, M.J., 2004a. The serial-parallel dilemma: a case study in a linkage of theory and method. Psychonomic Bulletin and Review 11, 391–418.

Townsend, J.T., Wenger, M.J., 2004b. A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. Psychological Review 111, 1003–1035.

Townsend, J.T., Yang, H., Burns, D.M., 2011. Experimental discrimination of the world's simplest and most antipodal models: the parallel-serial issue. In: Colonius, Hans, Dzhafarov, Ehtibar (Eds.), Descriptive and Normative Approaches to Human Behavior in the Advanced Series on Mathematical Psychology. World Scientific, Singapore.

Usher, M., McClelland, J.L., 2003. The time course of perceptual choice: the leaky, competing accumulator model. Psychological Review 108, 550–592.

Van Zandt, T., Ratcliff, R., 1995. Statistical mimicking of reaction time distributions: mixtures and parameter variability. Psychonomic Bulletin and Review 2, 20–54.

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

Wenger, M.J., Ingvalson, E.M., 2003. Preserving informational separability and violating decisional separability in facial perception and recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition 29, 1106–1118.

Wenger, M.J., Townsend, J.T., 2001. Faces as Gestalt stimuli: process characteristics. In: Wenger, M.J., Townsend, J.T. (Eds.), Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges. Erlbaum, Mahwah, NJ, pp. 229–284.