

Documents to U (MNU/MNUG)
University of Minnesota – Interlibrary Loan Lending
*OCLC MNU * RLG MNUG * NUC MnU*

15 Andersen Library, 222 21st Ave. S., University of Minnesota
Minneapolis, MN 55455-0439 USA
Phone: 612-624-4388, Fax: 612-624-4522, Ariel: 160.94.20.178, e-mail: docstou@tc.umn.edu

This article is delivered directly from the humanities, social & educational sciences, general sciences, mathematics, journalism and engineering collections of the University of Minnesota.
Thank you for using our service.

If you have problems with delivery, please contact us within 48 hours.

Notice: This material may be protected by copyright law. (Title 17 U.S. Code)

MNU

MNU

MNU

University of Minnesota - Interlibrary Loan Lending

Documents to U

15 Andersen Library, 222 21st Ave. S., University of Minnesota

Minneapolis, MN 55455-0439, USA

Phone 612-624-4388 • Fax: 612-624-4522 • Ariel 160.94.20.178 • email docstou@umn.edu

Web: www.minitex.umn.edu/dtu

If you had problems with this delivery, please reply within 48 hours:

- 1) Return this cover letter via Ariel or Fax
- 2) Indicate the problem
- 3) Indicate date received.

OCLC MNU RLIN MNUG

Notice: This material may be protected by copyright law. (Title 17 U.S. Code)

Thank you for using MNU – Thank you for using MNU – Thank you for using MNU

Trans. #:376119



ILL : 46497105



Article

Location : Wilson Library BF1.A4 A38x

Borrower : IUL

Date : 9/29/2008 01:06:31 PM

*MNU,DLM,WQM

AUTHOR :

TITLE : Advances in psychology.

IMPRINT : Amsterdam ; New York ; North-Holland Pub

Article : Townsend, J. T., & Evans, R.; A systems approach to parallel-serial testability and visual processing
Volume : 11 Issue : Month : Year : 1983 Pages : 166-191

ODYSSEY

FAX :

Copyright Compliance : CCL

Patron : Silbert, Noah

System ID : OCLC 5253350

NOTES :

Exempt

MaxCost: 30IFM

Loaned To:

IUL - Indiana University Libraries

DDS

1320 E.10th, #E065

Bloomington, IN 47405 USA

Loaned From : Documents to U : 15 Andersen Library, Univ. of MN 222 – 21st Ave S. Minneapolis, MN 55455

University of Minnesota Interlibrary Loan

ODYSSEY REQUEST

IR

A Systems Approach to Parallel-Serial Testability and Visual Feature Processing

JAMES T. TOWNSEND and RONALD EVANS

Introduction

In psychology, as well as other disciplines, we make theoretical advances through the use of scientific methodology. This methodology involves testing hypotheses about theories in experimental situations. When enough conflicting evidence or when a better theory surfaces, the old theory is typically replaced or modified. In many instances competing theories evolve simultaneously. In fact, it seems that in the social sciences, where so many hypotheses may appear equally viable from the outset, comparative testing of alternative theories may be the most effective path of progress. However, it happens rather frequently that although qualitatively different, competing theories are mathematically equivalent within certain experimental paradigms. This chapter deals with such an instance within the arena of mental processes.

An overview of this chapter might contain such key terms as 'parallel-serial processing', 'systems theory', and 'equivalence mappings'. The emphasis of the chapter will be placed on modeling and certain types of psychological systems within the context of systems theory, and testing classes of models of these systems.

Our investigations on parallel versus serial processing, and other modeling issues over the past fifteen years were neither motivated by nor conducted within a strict systems theory framework. Nevertheless there are a number of aspects of the work in systems theory, particularly the formal structures, that we believe may have significant epistemological value for scientists in the behavioral sciences. We thought it might therefore be interesting to attempt to place some of our recent results in the framework provided by systems theory, providing some of our views concerning its advantages and caveats concerning its proper implementation along the way.

In Section 1 a short history of the speed and timing of mental processing, beginning with Donders' Method of Subtraction and leading into the display and memory scanning experiments of the 1960's is presented. The relationship of the visual display and memory scanning experiments to the parallel-serial equivalence problem is developed at this time.

Section 2 gives a succinct presentation of systems theory and its application to the parallel-serial equivalence problem. Its potential use in formalizing a structure within which equivalence problem of models in general may be brought to the surface, is discussed. We will also provide a few remarks of a philosophical nature in this section.

The third Section presents a definition of parallel and serial models within the context of what we may call 'continuous time, discrete state stochastic systems theory'. We proceed at this particular point in the chapter to provide a fundamental distinction between parallel and serial processing and to suggest a promising mode of empirical attack on the problem.

The fourth and last Section of the chapter develops procedures using the ideas from Section 3 to test the hypothesis of a strict serial system in a distribution-free fashion.

These procedures are implemented in a whole-report accuracy experiment that records second guessing.

Basically, the distribution-free techniques are viewed as being a first order 'filter' for models while statistical fits of specific distributional models would offer a finer level of evaluation. If the fundamental qualitative (i.e., nonparametric and distribution-free) properties are not found in the data, no further tests should be carried out. But, if the predicted coarse constraints are satisfied, then specific models may be probed in an attempt to further narrow the class of acceptable theoretical explanations. The overall strategy is thus reminiscent of a series of sieves, each possessing progressively smaller openings; we will therefore refer to the general technique as a "*progressive sieve strategy*".

The general issue of serial versus parallel processing as well as certain other matters are discussed thoroughly in Townsend and Ashby (in press). The present treatment necessarily emphasizes a single disparity between serial and parallel systems. But, the systems interpretation as well as the distribution-free and statistical tests and the experimental application are new.

1. Mental processing and reaction time studies

Hypotheses about the brain and mental processing have progressed through time mainly by borrowing ideas from the technology of the current era. Early drawings of the brain by Descartes contain mechanical levers and pulleys acknowledging the early industrial era. By the 19th century, physiologists were aware of the electrical impulses traveling throughout the human body that transmitted messages for sensation and movement. Two prominent physiologists of that era, Müller and Helmholtz (the latter also being famous for his work in physics), had carried out experiments circa 1850 attempting to measure the velocity of the nerve impulses within the body. The hypothesis was that a nerve could be stimulated at a particular place and the subsequent reaction time for the twitch of a muscle connected to that nerve could be measured. The nerve was stimulated again, at some distance d from the original point of stimulation. The reaction time was measured again and the difference in the two reaction times would be the time the impulse took to travel the distance d .

Another prominent physiologist at that time, F. C. Donders (1969), took note of these experiments and wrote accordingly in 1868. Donders surmised that in an analogous fashion to Müller and Helmholtz, he could measure the time the mind took to execute various mental processes. By creating a succession of mental tasks to be undertaken and measuring reaction time for completion of these tasks, Donders felt he could measure the time for a specific task of interest by recording reaction time with the specific task included in the task set and then recording reaction again with the specific task taken out or subtracted from the task set. The time for completion of the specific mental task would be the difference in the two reaction times. The experimental method took the name *Donders' Method of Subtraction*.

Donders' methodology for measuring mental processes could be considered an attempt to systemize the timing of mental events for cognitive functions. Because of the importance of this work to reaction time experiments in general, we shall present an example of Dondersian procedure. One such experiment performed by Donders contained two control conditions, C1 and C2, where C1 measured the speed of the subject's button pressing response with the left hand to the onset of a light, and condition C2 measured the speed of the subject's button pressing response with the right hand to the onset of a light. An additive stage condition, C3, consisted of a presentation of one of two colors of light, upon which the subject responded by pressing a button with the left or right hand, the specific hand designating the fact that the subject observed a specific color light.

After running a number of trials the mean reaction time for each condition was computed. Denote these mean reaction times as $rt(C1)$ for mean reaction time in condition 1, $rt(C2)$ for mean reaction time in condition 2, $re(C3L)$ for mean reaction time in condition 3 given the response was made with the left hand, and $re(C3R)$ for mean reaction time in condition 3 given the response was made with the right hand. According to Donders' hypothesis

$rt(C3L)-rt(C1)$ and $rt(C3R)-rt(C2)$,

were both measurements of the mental processing time needed to distinguish the color of light.

A key assumption implicit in Donders' experiments was the presumed seriality of the processes, that is, that processes were carried out one after another with no temporal overlap. Initially this appeared to be a logical assumption for the types of tasks Donders used in his experiments, but, as in most experimental paradigms, strict seriality is not necessarily a true assumption when considered in the context of other processing designs.

Donders' work was taken note of by Wundt, who had opened what was later to be considered the world's first psychological laboratory at Leipzig in 1879. With the opening of his laboratory, Wundt and his students began conducting two basic types of experiments involving reaction time. These were:

- (1) simple reaction time experiments,
- (2) reaction time after subjective distinction of the stimulus or impression.

The 1st type involved a response after detection of any stimulus and the second type consisted of a response only after the subject had subjectively identified the stimulus. Thus the second type of experiment was an attempt to include the time necessary for the mental process of identification.

One of Wundt's students, Cattell (1886), continued Wundt's experiments with the purpose of devising a more objective means of measuring mental operations. Cattell was aware of Donders' work and criticized it on the premise that extraneous variables were involved in some of Donders' reaction time experiments. Cattell's specific criticism was that Donders' experiments usually involved two or more responses to differentiate identification of the various stimuli. Cattell claimed that this type of experiment involved what he termed "will-time" which was the time delay involved with associating the stimulus to its unique response.

Cattell (1886) carried out three types of experiments to identify the various components he thought were involved in the identification of a stimulus. These types were:

- (1) simple reaction time experiments to visual, auditory, and tactile stimuli,
- (2) reaction time experiments using the three modes above involving presentation of one of a pair of stimuli when a response was made after identifying one of the two stimuli but not after the other,
- (3) Reaction time experiments using the three modes above involving a multitude of stimuli presented one at a time in a random order with a unique response for the identification of each stimulus.

In (1) from above, Cattell also varied attention by instructing the subject to concentrate in one condition, use normal attention in another condition, and mentally add numbers in a third condition while awaiting the stimulus. From the differences in reaction time on the three treatments he concluded that simple reaction time tasks did not involve cortical function. From this he hypothesized that perception and willing were not factors of simple reaction time.

Results from the three classes of experiments also yielded the finding that $rt(1) < rt(2) < rt(3)$ in general for the same mode of sensation where $rt(i)$ is defined as the mean

me for each condition was com-
an reaction time in condition 1,
mean reaction time in condition 3
C3R) for mean reaction time in
t hand. According to Donders'

needed to distinguish the color

as the presumed seriality of the
fter another with no temporal
for the types of tasks Donders
paradigms, strict seriality is not
text of other processing designs.
d opened what was later to be
ipzig in 1879. With the opening
ting two basic types of experi-

ression.

stimulus and the second type con-
y identified the stimulus. Thus
he time necessary for the men-

Wundt's experiments with the
ng mental operations. Cattell
mise that extraneous variables
ents. Cattell's specific criticism
more responses to differentiate
his type of experiment involved
olved with associating the sti-

o identify the various compo-
stimulus. These types were:

e stimuli,

ng presentation of one of a pair of
muli but not after the other,
ng a multitude of stimuli presented
ntification of each stimulus.

ing the subject to concentrate
n, and mentally add numbers
differences in reaction time on
e tasks did not involve cortical
d willing were not factors of

added the finding that $rt(1) <$
ere $rt(i)$ is defined as the mean

reaction time for an experiment of type 'i'. Cattell interpreted the value for $rt(2)-rt(1)$ as being the perception time for a particular stimulus and $rt(3)-rt(2)$ as the will-time for choice of response associated with the correct stimulus.

Experiments of type (2) did give a refinement of a class of Donders' experiments under a restricted environment in that the temporal factor introduced by the will-time was removed. Type (2) experiments were not without fault, though. Cattell noted that there were trials on which the subject responded to the incorrect stimulus, i.e., false alarms. He did not deal with these trials, though, thus allowing the subject to set up some sort of probabilistic criterion resembling a threshold in signal detectability theory. Under this type of decision rule the subject could react when he or she was reasonably, but not entirely sure of the stimulus presented. Thus, it may have been possible for the subject to decrease reaction time simply because there was no penalty for a false alarm.

Research concerning attention and mental processing became secondary to related areas such as intelligence testing in the early 20th century. In spite of this shift in interest there were some experiments using variations of Donders' Subtraction Method being carried out. Hylan (1903) wrote a paper concerning the distribution of attention. Hylan had conducted some ingenious experiments involving reaction time in identifying various numbers of similar and dissimilar letters and objects presented simultaneously. From these experiments Hylan reached the conclusion that attention could not be distributed over objects, i.e., parallel processing could not take place.

His reasoning behind this conclusion, though, seemed to be tied in with a problem of semantics and definition as opposed to any true differentiation between parallel and serial processing. Hylan's reasoning from his reaction time studies was that distribution of attention, i.e., parallel processing, takes place only when conscious plurality has become conscious unity, that is elements are perceived as a whole rather than individual parts.

Hylan made an analogy between cognitive processing and electrical activity in the brain. His conjecture was that serial processing could be represented by spatially and temporally distinct electrical activity in the brain, i.e., the processing of different items took place in spatially distinct regions in the brain and that the serial processing signified a temporal distinction also with respect to the electrical activity. When the electrical activity took place in spatially distinct locations but occurred simultaneously, Hylan's interpretation was that this electrical activity now represented conscious unity or a Gestalt processing as opposed to 'parallel' processing of distinct items.

The reaction time experiment was not widely used again until the 1950's and 1960's where this time researchers found it useful for studying the cognitive processes of pattern recognition (e.g., Hick, 1952) display and memory search (e.g., Atkinson, Holmgren, & Juola, 1969; Sternberg, 1968, 1967), and multidimensional or multimodal pattern discrimination (e.g., Egeth, 1966). Of particular interest in such reaction time experiments were:

- (1) the assumptions concerning reaction time distribution,
- (2) the possible additivity of mean reaction time,
- and
- (3) processing independence across states and the stochastic latency mechanisms of reaction time in general (McGill & Gibbon, 1965; Falmagne & Theios, 1969; Sternberg, 1969).

Shortly thereafter, the expression of psychological mechanisms in terms of stochastic latency models and problems involved in testing alternative psychological assumptions came under investigation (e.g., Townsend, 1971, 1974; Taylor, 1976; Theios, 1973; Pachella, 1974; Kantowitz, 1974).

Many researchers (Sternberg, 1966, 1967; Neisser, Novick, & Lazar, 1963; Atkinson, Holmgren, & Juola, 1969) found that the time it took to search through a list of items for a target was linearly related to the length of the list, which promoted the notion of

serial scanning (see below or Townsend, 1974, for alternative interpretations). Conflicting results were reported concerning the termination of the scanning process, though. Sternberg (1966) and Atkinson et al. (1969) reported that the slope for mean reaction time as a function of list length did not vary for positive searches (where the target was in the list) and negative searches (where the target was not contained in the list). Since a negative search was logically interpreted as an exhaustive search of the item list, this finding seemed to mean that positive searches were exhaustive also, or that processing continued after the target was found.

Sternberg (1967), in contrast to the above finding, reported that the slope of the mean reaction time as a function of list length for positive searches was one half the slope of negative searches. This result appeared to support the hypothesis of a self-terminating serial model of processing via the logic that given random placement of a target in a list the change in scanning time would only increase half as much, on the average, for a positive self-terminating search as opposed to a negative, exhaustive search with respect to changes in the list length. (See Townsend, 1974, and Townsend & Ashby, in press, for a more definitive presentation of the self-terminating vs exhaustive processing issue.)

In addition to searching for a single target, Sternberg (1967) and Neisser et al. (1963) had subjects memorize multiple target items and then search through a list for any one of the possible targets. In this type of experiment, a logical processing model would involve not only the scanning process of the list but multiple matches with each item of the list to all of the target items in memory.

Neisser et al. (1963) found that with practice subjects performed as quickly with multiple targets as with only one target in memory. From this data he concluded that the matching process was actually a parallel process. Actually, parallel processing does not inevitably produce flat reaction time curves, nor does it necessarily imply unlimited capacity and/or correlated processing times (see, e.g., Townsend, 1974). In any event, we began to find in the late 1960's and early 1970's that model equivalences and limitations in observability provoked difficulties in testing latency based information processing theories that were akin to those discovered in mathematical learning theory (e.g., Greeno & Steiner, 1964).

For instance, although the support of serial processing models based upon linear mean reaction time data and slope with respect to change in list length seemed intuitively appealing, this approach proved to be methodologically problematic. Townsend (1971, 1972, 1974, 1976b), Vorberg (1977), and Anderson (1976) derived functional equations for classes of parallel and serial processing models which yielded mathematically equivalent models as well as equivalent mean reaction time predictions.

Many of the results reported above were primarily 'negative' in the sense of revealing intuitively plausible alternative models which could make the same predictions as a previously favored model. However, as work progressed, possibilities for testing parallelity versus seriality began to surface. For instance, Townsend (1976a) derived an experimental situation and Townsend and Snodgrass (1974) tested this situation where parallel and serial processing could be distinguished from each other at the level of the mean reaction time curve given the following two conditions:

- (1) the rate for a positive match between a target and item was different from the rate for a negative match between target and item,
- (2) processing is self-terminating in 'target present' conditions.

Snodgrass and Townsend (1980) employed similar tests in an analogous, if somewhat more complex paradigm.

Although we feel the above and certain other techniques offer substantial hope for experimentally testing the two types of processing, the differentiating aspects on which

the present work is based appears to be among the most fundamental distinctions between parallel and serial models (Townsend & Ashby, in press).

Before applying systems theory to the parallel-serial problem in this chapter we provide an introduction to systems theory. The introduction given in Section 2 has been written with the intent of relating some basic ideas about systems theory with references for those possessing more intense interest. The mathematical framework in Section 2 is kept to a minimum; basically only what is necessary for specific discussion of the parallel-serial problem in Section 3 is developed.

2. Introduction to system concepts

To introduce the notion of a system, whether it be a psychological system or a highly structured electronic system (or whatever), we begin with a general definition. A system can always be characterized by an input(s), a process or function applied to the input(s), and output(s) that is (are) the result of this transformation. Fig. 1 is a pictorial representation of this notion.

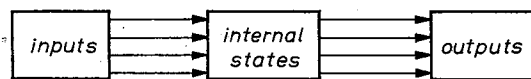


Fig. 1: Illustrating the general and primitive notion of a processing system.

With this simple notion of a system, we can broadly define *general systems theory* as a metatheory devoted to the study of general and abstract properties of real world systems. Within it one may formalize the construction of systems in terms of models which have specifically defined characteristics. This definition may appear quite general in nature, and, in fact, is meant to be. Klir's (1972) book contains a variety of notions concerning the uses and definitions of systems theory as proposed by the authors of the various chapters in the book. His introduction of the presentations of systems theory in the various chapters proceeds as follows:

General systems theory is considered as a formal theory (Mesarovic, Wymore), a methodology (Ashby, Klir), a way of thinking (Bertalanffy, Churman), a way of looking at the world (Weinberg), as search for an optimal simplification (Ashby, Weinberg), an educational tool (Boulding, Klir, Weinberg), a metalanguage (Lofgren), or prospectively, a profession (Klir).

Thus, even various authors within specific disciplines hold different ideas about the uses and meaning of systems theory. As will be brought out in this chapter, we shall see what use information processing psychology may have for systems theory.

The advent of modern day systems theory in the 1930's is generally attributed to Bertalanffy (1932). The idea of a systems approach to problems is not new, and is in fact related to the Aristotelian view of the world with holistic and teleological notions. In reference to this point, we find that the underlying principle in systems theory is contained in Aristotle's statement, "The whole is more than the sum of its parts", also the main thrust of Gestalt psychology.

We find mention of holistic and Gestaltists' views in much of the literature previous to Bertalanffy's theorizing but the reality of the situation was that classical science was hampered in dealing with multiple interactions or relationships. An example by Weaver (1948) stated that:

„Classical science was concerned with one-way causality or relations between two variables, such as the attraction of the sun and a planet, but even the three body problem of mechanics permits no closed solution by analytical methods of classical mechanics.”

Newtonian physics embodies the important characteristics of classical science. At some risk of oversimplification it may be claimed that this method consisted of the reductionist philosophy, i.e., breaking the system down into individual components with the important assumption being the independence of these components. Solution for the independent parts were found and thus the problem was considered solved.

It appears intuitive to suggest that computers, for the most part, had the greatest impact on the study and developmental of systems theory. The iterative potential of the computer is useful for solving the large number of matrix equations existing in most closed form systems problems. In those cases where closed form solutions are not possible, open form solutions may still be arrived at through iterative approximation where the computer again allows the flexibility to include a large system of equations. Properties of many systems such as continuous time also lend themselves to methods of numerical approximation by iterative techniques.

Mesarovic (1968) proposed a methodology for applying systems theory to biological problems. A summary of this approach for psychological systems is shown in the flowchart of Fig. 2.

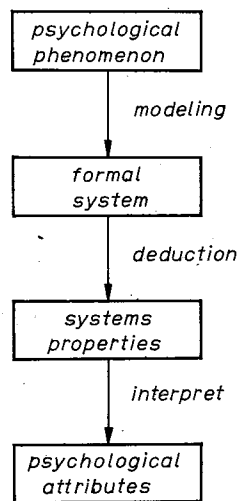


Fig. 2: Schematic of a systems modeling approach (from Mesarovic, 1968).

As shown in the flowchart, after the initial characterization of the formal system, specific models applying predetermined assumptions may be devised and then tested.

Many problems will not require the entire framework provided by the general systems theory as will be the case in attacking the parallel-serial equivalence problem. To more fully understand what follows, though, we will introduce here a more formalized notion of a system.

Formalization of a system

We wish to introduce the notion of a system which shall be denoted as S . For a more detailed introduction interested readers may refer to Kalman (1965, 1969) or Arbib (1965).

relations between two variables, such as the problem of mechanics permits no closed

of classical science. At some point the method consisted of the reductionist approach, breaking down the system into its component parts with the important components with the important components. Solution for the independent components was never solved.

most part, had the greatest importance. The iterative potential of the iterative method existing in most closed form solutions are not possible, iterative approximation where the iterative method of solving the system of equations. Properties of iterative methods themselves to methods of numerical

ing systems theory to biological systems is shown in the flow-

The concepts and notation for the following definitions were chosen from the references above. For the purpose of brevity, though, only the more important concepts were chosen for discussion in this section.

S consists of a 6-tuple $\langle T, U, X, Y, \varphi, \eta \rangle$ defined as follows:

$T \equiv \langle \text{Set of all times or time environment within which S operates} \rangle$

$U \equiv \langle \text{Input space or set of all possible inputs to S} \rangle$

$X \equiv \langle \text{State space or set pertaining to all possible internal conditions of S} \rangle$

$Y \equiv \langle \text{Output space or set of all possible outputs from S} \rangle$.

And φ and η are mappings defined as follows:

$\varphi: T \times T \times X \times U \rightarrow X$ where φ is the state-transition function, and

$\eta: T \times X \rightarrow Y$ where η is the readout map.

The input to the system beginning at time t_1 and ending at the instant before time t_2 is denoted as $w(t_1, t_2)$. The bracket expresses the convention that the t_1 instant is included in the temporal interval plus all instants greater than t_1 and less than t_2 (but not including t_2).

The state-transition function for a particular set of values can then be expressed as follows:

$$\varphi(t_2; t_1, x(t_1), w[t_1, t_2]) = x(t_2) \in X \text{ for } t_1 < t_2. \quad (3A)$$

This equation expresses the fact that knowledge of the state at time t_1 and knowledge of the admissible input from time t_1 until the instant before time t_2 is sufficient information to determine $x(t_2)$, the state of the system at time t_2 . Implications about the procession of time from past to present and the causal relationship of the system are borne out by the state-transition function, i.e., the future may be determined by knowledge of the past.

At this point it is worth emphasizing that the state-transition function presented thus far is for a deterministic system, that is, there is a one-to-one or many-to-one mapping from past inputs and states to future state. In the next section we will redefine the state-transition function for a stochastic system where a probability distribution exists for each possible transition.

The final mapping of the 6-tuple we need to consider is the readout map which takes on the following form for a set of specific values:

$$\eta(t_1, x(t_1)) = y(t_1) \in Y. \quad (3B)$$

This function provides a deterministic mapping from a state at a particular time to the output of the system for that time. Here again, many psychological systems may be represented better by a stochastic readout map. This consideration will also be taken up in the next section. Combining equations 3A and 3B we have the following result:

$$\eta(t_2, \varphi(t_2; t_1, x(t_1), w[t_1, t_2])) = y(t_2) \text{ for } t_2 > t_1. \quad (3C)$$

Therefore we have an equation which provides us with the output of the system at time t_2 given we know the state of the system at some time prior to t_2 and that we know the input sequence between times t_1 and t_2 .

Systems operate in real life situations. When we wish to study a system, however, it is customary to observe the system within a carefully constructed and constrained framework. The present framework is an experiment. The importance of the experimental framework cannot be stressed enough. For a systems framework we define an experiment as follows:

Experiment

An experiment, E , consists of 3-tuple (U, Y, T) where

U = the set of allowable inputs,
 Y = the admissible set of outputs,
 T = the space of time within which the experiment operates.

For instance in psychology, two types of experiments often used to test information processing characteristics are reaction time experiments and accuracy experiments. Although it is of course sometimes possible to combine them in the same paradigm, if more than one response is required, the later response times, may not be meaningfully related to the psychological process under study. The set of inputs, their timing and their relation to the set of outputs tend to differ for the two varieties of experiment. As more systems terms are defined in this and later sections, the reader should take note of the importance of the role the systems experiment actually plays in defining what can and cannot be measured or construed from a system.

It is essential that a system be compatible with an experiment if the system is to operate within a specific experimental framework. As we have defined it, if $U_1, Y_1,$ and T_1 are the input, output, and time space of a system, L , and $U_k, Y_k,$ and T_k are the respective sets for an experiment, K , the system L and experiment K are compatible if and only if

$$U_1 \wedge U_k \neq 0, Y_1 \wedge Y_k \neq 0, \text{ and } T_1 \wedge T_k \neq 0.$$

As an example, consider an experiment which uses analog information. If the system we wish to run the experiment on handles only discrete information, the experiment and the system are not compatible.

As noted in the flowchart diagram of Fig. 2, and also well known in the area of psychology, it is convenient to construct a model, M , of a system which is compatible with particular sets of experiments. An informal definition of a model is as follows¹:

Model definition

A model, M , consists of a 6-tuple $\langle T, U, X, Y, \varphi, \eta \rangle$, where each of the six sets may be defined as being some subset of its respective component in the system, S . These subsets are defined by specific assumptions and rules chosen by the investigator.

Examples of two common assumptions or rules applied in the psychological literature are:

- (1) Distribution assumptions used in reaction time studies which may eliminate subsets of possible state or output variables.
- (2) Optimal guessing or partitioning strategies used in decision theory (one of these which will be used in the model of the whole-report experiment at the end of this chapter).

In many, perhaps most, engineering applications, one attempts to design the system with certain inputs, state mappings, and desired outputs in mind. In psychology and in 'black-box-engineering and computer science analysis' situations the precise φ and η mappings

¹ In some contexts it is helpful to emphasize the distinction between an abstract characterization and the real-world, nuts and bolts system itself. In such cases, we have referred to the real-world object as the "system" and the abstract characterization as the "model" of the system (Townsend & Ashby, in press). In much writing in the systems science area, the term "system" is used to refer to the abstract characterization. It will be convenient, and do no harm, to follow this convention here and to conceive of a model as a specified system.

are unknown and it is the research objective to ascertain their nature. The property of *system observability* thus becomes an issue. A standard systems definition is as follows:

System observability

A system is observable in an experiment if and only if the initial state, $x(t_0)$, can be determined from the input-output pairs of the experiment.

This definition is much more powerful than it initially appears to be simply because t_0 can be any value, so in principle any state must be determinable by observing the input-output pairs across time. This definition as well as the one below is perhaps most readily suited to 'experiments' where for every input and internal state there is an output. Many practical engineering situations are of that sort. Many empirical paradigms in psychology are not so fortunate.

In fact in typical cognitive experiments, after a stimulus input, the information processing system goes through an entire course of state changes before emitting a single response. The investigator must ordinarily employ an aggregate of stimulus and response trials using statistics on reaction time accuracy and/or other dependent variables to explore the underlying systemic structure of the organism.

Another important notion in systems theory is that of *controllability*. Informally, the property implies that one can force the system into any of its possible states by proper manipulation of the inputs. More formally we have:

System controllability

A system is controllable in an experiment if and only if for any $x \in X$, $\varphi(t_2; t_1, x(t_1), w[t_1, t_2]) = \theta$ for some $t_2 \in T$ where $\theta \in X$ is a predesignated state.

This definition simply assures us that we can force the system into the particular domain of our interests. Although little applied at least in a formal manner to psychological variables, it will undoubtedly come to be of considerable import in such areas as man-machine integration, motor control theory, optimal continuous time detection jobs, and so on.

For most psychological research it is presently too much to hope that data from our experiments will provide us with the actual mappings discussed for all admissible values of inputs, outputs, and states. Observability and controllability over the entire system is also most likely beyond our reach. What we may be able to obtain, though, through correct experimental procedures, is observability and controllability over some subset of the state space which might be of specific interest for delineating classes of theoretical explanations which are capable of predicting and clarifying a given data base.

In a sense, modern psychology has always worked somewhat along the lines of the 'black box' technique in order to come to an understanding of the way the mind works, only it was not put in those terms. Furthermore, the propensity of those interested in such matters to borrow ideas from the physical sciences and objects goes back in time about as far as one wishes to look. The famous portrayal by Descartes of man as an automaton, characterized by the intricate clockwork figurines of the time, with the addition of a soul which interacted with the body-automaton through the pineal gland, captures the spirit of this inclination. Perhaps only the radical Skinnerians (via Skinner's early precepts) can claim to have avoided this tendency.

One of the propitious aspects of modern systems theory is its very generality. As a case in point, a frequently heard criticism of the information processing approach is that it leans too hard on the digital computer analogy. It is thus said to be too narrow and imitative of 'dumb' computers, and possesses certain properties not because that is how

they really are, but because that is how computers are presently being built. The general concepts of systems theory provide a matrix within which a tremendous variety of already realized systems reside, but more importantly, within which new systemic architecture may be invented that is most appropriate to a specific class of environmental and psychological situations.

As an example, not only the theory of automata (the theoretical underpinning of the digital computer) is a specific type of system, so also are continuous time linear systems, which immediately offer the potential for continuous flow or cascade models of processing (e.g., Hoffman, 1978; Ericksen & Schultz, 1979; McClelland, 1979; Townsend & Ashby, in press, CH 12) as well as parallel and hybrid types of processing. Even the currently attractive holographic models of perception and memory may be studied within the confines of system theory.

In the realm of behavioral psychology and other black box disciplines, the problem of system and model equivalence is critical. Although it is possible to define equivalence at the level of identical state and output mappings, such is of little interest here. More pertinent is what we might call "behavior equivalence"; however because we are working only with this type, we will drop the "behavior".

System-model equivalence

The models, M1 and M2, of two systems are equivalent in an experiment $E = [U, Y, T]$ if and only if for equivalent inputs, $w_{m1}[t_1, t_2] = w_{m2}[t_1, t_2]$, the outputs, $Y_{m1}(t_2)$ and $Y_{m2}(t_2)$, are equivalent (where $w_{m1}, w_{m2} \in U, Y_{m1}, Y_{m2} \in Y$, and $t_1, t_2 \in T$).

One might define system equivalence in such a way that the input state and output sets are not the same, yet the behavior of the two systems is entirely 'isomorphic' in the two models, but there is little point in doing so here. Notice that the subsystems and their hook-ups inside the overall system might be quite different in their 'anatomy' or 'wiring' yet perform in exactly the same fashion. Examples abound in linear systems theory and automata theory but are beyond the present space latitudes (but see Booth, 1967 in the case of theory of automata and DiStefano, Stubberud, & Williams, 1967 in the case of linear control systems).

As noted earlier, we need not, indeed, must not slavishly copy any particular class of systems. Rather we must carefully borrow those concepts and techniques that seem most helpful and germane to the psychological quest. A case of where we need to develop more systems theory that is uniquely apposite to psychology is continuous time stochastic systems theory. Although literature on this topic exists in the engineering sciences, the reader will soon find that the theories are ill suited to psychological theorizing. For instance, in most engineering theory there is little concern with variation in processing time of a system that is inherent in the system itself. This is because in most applications, the system may be considered to be deterministic and all of the variability is placed in the signal. However, the view of most psychologists would probably be that in human mental processing the mental processes themselves are inherently stochastic in terms of how long they require to perform a given task, as well as perhaps the particular output they emit at a given point in time.

Actually, in many of the contexts where applied mathematicians talk about stochastic systems theory, there is almost no consideration of duration at all. The quality of description of the signal itself is of paramount interest. Now, in theory, one should be able to calculate the time needed to finish a task (or more precisely the probability distribution on that duration) from the type of description typically offered, but in fact it may be nearly impossible in realistic circumstances. In the few applications of linear systems theory (to take an example) to cognitive psychology, one usually ignores variance in processing

time.
mula
a thr
sourc
both
syste
ed in
syste
Th
discr
put &
syste
are b
to th
struc
of th
appr
cussi

3. S

We a
The
secti
We
one a
syste
proc
Th
infor
psycl
proc
digit
case,
that
proc
each
plete
nuou
Fo
thoug
one

(1) co
(2) pa
(3) co

We v
vario
are d
proce
whicl

12 Ge

sently being built. The general tremendous variety of already rich new systemic architecture of environmental and psycho-

oretical underpinning of the continuous time linear systems, or cascade models of processing (McClelland, 1979; Townsend & Sengco, 1979). Even the current memory may be studied within

other disciplines, the problem of possible to define equivalence at a level of little interest here. More interesting however because we are working

Experiment $E = [U, Y, T]$ if and only if E_1 and E_2 are equivalent (where

the input state and output sets are entirely 'isomorphic' in the two systems, that the subsystems and their 'anatomy' or 'wiring' are identical in linear systems theory and practice (but see Booth, 1967 in the case of Williams, 1967 in the case of

any copy any particular class of models and techniques that seem more useful here we need to develop more models of continuous time stochastic systems. In engineering sciences, the reader interested in theorizing. For instance, in the case of processing time of a system or in applications, the system may be placed in the signal. However, in human mental processing the terms of how long they require to produce an output they emit at a given

mathematicians talk about stochastic models at all. The quality of descriptive theory, one should be able to specify the probability distribution function, but in fact it may be nearly identical to that of linear systems theory (to ignore variance in processing

time. A recent linear systems model adds a random base activation to the activation accumulated in a set of linear integrators (McClelland, 1979). A response is determined when a threshold criterion is reached and the random amplitude component thus instals a source of reaction time variance. An example of a psychological model incorporating both a random decision threshold as well as randomly additive activation within a linear systems framework is that of Pacut (1981). It appears that the stochastic element contained in the describing differential equation of Pacut could be considered as inherent in the system itself.

There seem to be at least two potential ways out of this dilemma. One would be to use discrete time approximation. It is easier when dealing with an output for every input, to put a direct stochastic time element in such systems than it is with continuous time systems. Another course would be to develop theories that, like the engineering theories, are based on probabilism connected with the input-output train or sequence. In contrast to that discipline, however, the theorist would take the option of putting the stochastic structure directly into the state and output mappings, rather than in the stochastic nature of the signal. (Obviously there are cases such as signal detection where the engineering approach is the most natural.) In any event this may give some hint of the kinds of discussion that may fruitfully (we hope) take place within the systems framework.

3. Stochastic systems approach to the parallel-serial equivalence problem

We are now ready to consider modeling the human visual information processing system. The complexity of this system causes us to modify the system definitions from the last section.

We must first consider the stochastic nature of the visual processing system. Whether one argues that the stimulus input to the system is simply too complex or that the human system contains components of 'free will', we concede that, for the present, an information processing model of the human system must be probabilistic or stochastic in nature.

The second consideration for modeling purposes is to specify the state of a piece of information at any time. The type of information that the system must deal with in most psychological situations is a list of conceptually distinct items. In the context of the visual processing system, items may typically be thought of as letters of the alphabet, individual digits or numbers, or sometimes specially devised nonsense symbols. In the most general case, the processing of each item takes place on a continuous basis. By this, it is meant that the state of any particular item may vary from completely unprocessed to completely processed. The opposite end of the spectrum may be a type of all-or-none state space for each item. In this case the state of an item may take on only the 'binary value' of completely unprocessed or completely processed. Any partially discrete and partially continuous approach between these two ends of the spectrum may exist as well.

For the express purpose of modeling serial and parallel processing systems in Section 4, though, it will eventually serve our purpose to classify a piece of information as being in one of three possible categories of states:

- (1) completely unprocessed,
- (2) partially processed,
- (3) completely processed.

We will typically presume that items in category (2) assume discrete values representing various degrees of processing, but this is not necessary. Thus, the model of the system we are dealing with will ultimately be classified as a *discrete state continuous time stochastic processing system*. Given this type of system we wish to supply a few formal definitions which are applicable to the situation.

Definition of discrete state stochastic system

Let $S = \langle T, U, X, Y, P_p, P_q \rangle$ where $T, U, X,$ and Y are defined as before. P_p is a stochastic process on the state transitions such that

$$P_p[X(t_2) | t_2; t_1, X(t_1), w[t_1, t_2]]$$

is the probability that the system is in state $X(t_2)$ at time t_2 given the system state at time t_1 was $X(t_1)$ and the input in the interim was $w[t_1, t_2]$. Note that the process $P_p(-)$ is MARKOV in state and input since knowledge of previous state and input before t_1 is irrelevant given the state $X(t_1)$ and the input $w[t_1, t_2]$ is known. Similarly, P_q is a stochastic process on the output mappings in the sense that

$$P_q[Y(t) | t, X(t)]$$

gives the probability that the output is $Y(t)$ at time t where the system state is $X(t)$.

In some instances, given the state of the system, one might expect a specific or determined output. For a system that processes information, though, any state designating partial or no processing is capable of producing a distribution of outputs which is dependent upon the output space of the experiment. As an example, consider a multiple choice test. On many questions the system or person does not have complete knowledge that is necessary for correctly picking the correct output. The state of the system might be termed *partial knowledge* on some questions. Given the specific state, the system is able to eliminate some invalid alternatives and choose randomly or with some specified distribution from the remaining alternatives. The type of system that processes in such a manner is an important concept. This characteristic is desirable of most processing systems and the assumption is commonly used in most psychological models of information processing. We identify this type of system as an intelligent processing system and define it as follows.

Definition for an intelligent processing system <IPS> within a low accuracy systems experiment

An intelligent processing system may be defined according to two major assertions. First, the output of an intelligent processing system will always reflect at least the amount of information the system has obtained about the stimulus configuration. Second, the system must be stochastic, not deterministic.

As an example of such a system, consider a card player who is trying to guess which card is on top of a deck of cards. Initially, his chance of guessing correctly is one out of fifty-two. If the card player knows the card on top of the deck is a black card, it is only reasonable to assume that he will now narrow his choice of guesses down to all the black cards in the deck. Therefore, using the <IPS> assumption, we would now assess his chance of guessing the correct card as being one out of twenty-six. Thus, if the card player was questioned fifty-two times about the top card in the deck, we would expect him to be correct about twice. Therefore, on some trials or guesses, even though the processing system had only partial information about the state of affairs, he was able to exhibit an output (guess) which made it appear that the system had substantially more information about the top card in the deck (i.e., those guesses which were correct). If we look at all of the fifty guesses made by the card player we would probably find that all of the guesses were black cards. Therefore, all of the outputs reflected at least the amount of information that the state of the system had obtained about the stimulus input. Now suppose the card player is also given the information that the card on top of the deck is a queen. The <IPS> assumption would now allow us to assess the card player's chance of being correct

as before. P_p is a stochastic process

the system state at time t_1 was $X(t_1)$
(—) is MARKOV in state and input since
the state $X(t_1)$ and the input $w[t_1, t_2]$
ings in the sense that

system state is $X(t)$.

nt expect a specific or determined
gh, any state designating partial
n of outputs which is dependent
, consider a multiple choice test.
, complete knowledge that is ne-
e of the system might be termed
c state, the system is able to eli-
with some specified distribution
processes in such a manner is an
most processing systems and the
odels of information processing.
g system and define it as follows.

within a low accuracy

vo major assertions. First, the output
amount of information the system has
ust be stochastic, not deterministic.

ho is trying to guess which card is
g correctly is one out of fifty-two.
black card, it is only reasonable
down to all the black cards in the
would now assess his chance of
ix. Thus, if the card player was
eck, we would expect him to be
ses, even though the processing
affairs, he was able to exhibit an
d substantially more information
were correct). If we look at all of
bably find that all of the guesses
at least the amount of information
stimulus input. Now suppose the
n top of the deck is a queen. The
player's chance of being correct

at one-half (i.e., the card is black and it is a queen, therefore it must be either the queen of spades or the queen of clubs). Once again, if we asked the card player to make fifty guesses, his output (guesses) would either be the queen of spades or the queen of clubs. The output would, once again, reflect at least the amount of information the system had obtained about the input.

An important special case of the \langle IPS \rangle assumption is the notion of "sophisticated guessing" in recognition-confusion experiments (e.g., Broadbent, 1967; Townsend & Landon, in press a, b). The idea there is that when a noisy or fragmented stimulus pattern is received by an observer, he or she is able to use the partial information to pare down the list of potential alternative signals to a so-called "confusion set". The observer then employs some guessing strategy to select a response from that delimited set of possibilities.

The \langle IPS \rangle assumption is very simple, indeed, and we doubt that there are any human information processing models which do not use this assumption. The important point that we wish to emphasize, though, is that from a systems theory viewpoint, this assumption must be formulated and stated as one of the model's axioms. The topic concerning optimal use of the system state in choosing an output will also be dealt with in Section 4.

If we now assume we are dealing with an \langle IPS \rangle operating in an imperfect accuracy environment, we no longer have observability as defined in Section 2. What we do have might be termed statistical observability where for individual trials or realizations of an experiment we have no observability for many of the states. However, over many trials we may have statistical observability for some states which were not observable for individual trials by estimating guessing probabilities for such occurrences. Referring back to the card player's problem, we could statistically test the hypothesis that the card player knew the top card of the deck was a queen simply by observing the number of times the card player's guess was a queen and comparing this number to the random probability of choosing a queen, one out of thirteen. Once again, we point out that the idea of statistical observability is commonly used in most testing frameworks, but, it is necessary to classify the idea within the systems theory framework. We will deal with statistical observability later in this section, as well as in the tests used in Section 4.

At the present time, we wish to present definitions which lay the foundation for the systems approach to the parallel-serial study. Initially, we must have a way of defining the input to the system. This will be as follows:

Definition of an item or input component

An item or input component is assumed to be made up of features and can therefore be defined by a vector

$$s_i = [s_{i1}, s_{i2}, \dots, s_{in}] ,$$

where s_{ij} is the j th feature of item 'i'. Thus, the global stimulus display may be represented by a vector of vectors, each subvector designating an item by its features.

Given the input to the system is as such, we now present two typical types of experiments. Note in these two definitions that actually the form of the input to the system is described, and not the output. The wide class of outputs which may be obtained varies greatly with the paradigm.

Standard parallel-serial reaction time experiment for fixed n with perfect accuracy

A set of items is presented to the system simultaneously for a given duration. Sometimes the stimulus duration might last until a designated response is invoked and sometimes it might be preset. It is assumed

here that the items are processed until enough information is accumulated to make a correct response. In this type of experiment one may represent the input function to the processor as a vector

$$w(t) = [w_1(t), w_2(t), \dots, w_n(t)],$$

where $w_i(t)$ gives the time course of the *system response* to the actual stimulus input which is typically a square wave (as much as possible); that is, all items are *turned on or off* instantaneously and simultaneously by the apparatus used to present the stimuli.

In many situations, it may be reasonable to assume for practical purposes that $w(t)$ is a step function which lasts until requisite processing is complete so that we may consider the input as a constant set of items throughout processing.

Parallel-serial experiment for fixed n : high resolution and sharp masking

This experiment is analogous to the previous experiment except that a square wave (in time) mask comes on at some point which effectively obliterates the display. It is not necessary to assume the display 'image' in the head is erased, only that information processing capability is essentially brought to a halt.

We have purposefully left out the description of the task in the two definitions above. By defining the tasks to be performed we define the output set for the experiment which we wish to leave open for the moment. With these definitions in mind we turn to the specific problem of representing a *discrete state stochastic serial processing model* and a *discrete state stochastic parallel processing model*. As systems both of these models may be represented by the 6-tuple

$$S = [T, U, X, Y, P_p, P_q].$$

Our interest is to define both of these types of systems without any dependence upon the experimental framework. As mentioned earlier, the input, output, and time sets must prove to be consistent or compatible with those of the experimental framework. The stochastic output function, P_q , is dependent upon response alternatives, causing this function to be indirectly dependent upon the specific experiment. Eliminating these sets as possible sources for definitional characteristics of parallel and serial processing models, we are left with X , the state space set, and P_p , the stochastic process on state transitions. It is these two sets within the 6-tuple definition of a system which are the natural defining operators of parallel and serial processing systems. Before using these concepts, though, in a proper definition, a clarification of the way we intend to use the state space set and stochastic transition function is in order.

Discrete item, state space definition

Let X be the set of admissible states. Characterize X by a vector $\langle x_1, x_2, \dots, x_n \rangle$ where x_i denotes the proportion of item 'i' processed, $0 \leq x_i \leq 1$ and $1 \leq i \leq n$. The state of a display which is completely processed is then represented by the vector $\langle 1, 1, 1, \dots, 1 \rangle$ while the initial state of a display before processing is represented by the vector $\langle 0, 0, 0, \dots, 0 \rangle$.

It is a trivial matter, though, to map the states as they would be classified from the definition above to the three categories previously noted as (1) completely unprocessed, (2) partially processed, and (3) completely processed.

Given the definitions for the stochastic system, we now present qualifying definitions (axioms) for the serial and parallel systems.

accumulated to make a correct response.
to the processor as a vector

actual stimulus input which is typically
turned on or off instantaneously and simul-

cal purposes that $w(t)$ is a step function
consider the input as a constant set of

and sharp masking

that a square wave (in time) mask comes
necessary to assume the display 'image'
is essentially brought to a halt.

work in the two definitions above. By
the set for the experiment which we
keep in mind we turn to the specific
serial processing model and a *discrete*
form of these models may be repre-

without any dependence upon the
input, output, and time sets must
be the experimental framework. The
response alternatives, causing this
experiment. Eliminating these sets
parallel and serial processing models,
stochastic process on state transitions.
from which are the natural defining
more using these concepts, though,
tend to use the state space set and

for $\langle x_1, x_2, \dots, x_n \rangle$ where x_i denotes the
state of a display which is completely
while the initial state of a display before

could be classified from the defini-
s (1) completely unprocessed, (2)

now present qualifying definitions

Discrete state (in item length) serial stochastic processing system

A processing system is serial in nature if and only if the state vector, X , has the following restriction. For all $x \in X$ where $x = \langle x_1, x_2, \dots, x_n \rangle$

- (1)
- (A) $x = \langle 0, 0, 0, \dots, 0 \rangle, t = 0$
- (B) $x_i = 0$ or 1 for all 'i' ranging
between 1 and n excluding 'j'
where 'j' may be any one integer
between 1 and n (e.g.,
 $\langle 1, 1, .5, 0, 0, 0 \rangle$ for $0 < t$)

$$0 \leq x \leq 1 \text{ for all } 0 \leq t.$$

- (2)
- $$p(x(t_2) | t_2; t_1, x(t_1), w[t_1, t_2]) = 0$$
- if $x_i(t_2) < x_i(t_1), 1 \leq i \leq n, 0 \leq t_1 < t_2.$

This definition, which we later rename the *serial systems axiom*, is a result of the serial system only working on one item at a time. Therefore, if we observe the state of the system at any particular time, all items but one must be completed or not begun yet.

Part (2) simply states that information can only advance, never disintegrate or recede.

Discrete state (in item length) stochastic parallel processing system

A stochastic processing model is parallel in nature if and only if the following two propositions hold.

- (1)
- For all $x_i \in X$ where $x = \langle x_1, x_2, \dots, x_n \rangle$ either
- (A) $x = \langle 0, 0, 0, \dots, 0 \rangle, \text{ at } t = 0$
or
- (B) $0 \leq x_i \leq 1 \text{ for } 1 \leq i \leq n, 0 < t$

and
(2)

$$p(x(t_2) | t_2; t_1, x(t_1), \omega[t_1, t_2]) = 0$$

if $x_i(t_2) < x_i(t_1), 1 \leq i \leq n, 0 \leq t_1 < t_2.$

Part (1) of this definition sets up the acceptable states of processing, note that no states are ruled out after $t > 0$. Part 2 implies that the processing state of an item, viewed through $x_i(t)$, can only advance, as in a serial system.

Our knowledge of a systems experiment and a systems representation now allows us to describe a general algorithm for distinguishing parallel and serial processing models. Let

$U(S)$ = input set for an, as yet, undetermined type of processing system,
 $U(E)$ = input set of experiment E ,

$Y(S)$ = output set of the, as yet, undetermined type of processing system,
 $Y(E)$ = output set of experiment E,
 $X(s)$ = state space set of a serial processing model,
 $X(p)$ = state space set of a parallel processing model.

Now when we perform the experiment on the unknown processing system, we observe the input-output pairs $U(S) \wedge U(E)$ and $Y(S) \wedge Y(E)$. If we have observability in our experiment, we are able to determine a state space set, $X(S) \wedge X(E)$, resulting from the input-output pairs. We have the capability of distinguishing if the model is parallel or serial if

$X(S) \wedge X(E)$ is contained in $[X(s) \vee X(p)]$.

In other terms, we are saying that if we have a limited type of controllability, we may force the observed state space to be contained within the state space which distinguishes a parallel processor from a serial processor.

For clarity, we present an instance of a processing system which is given three items to process. Consider three possible states of processing for each item. Let these three states be denoted as:

- (i) zero processing state denoted as 0,
- (ii) partial processing state denoted as 1,
- (iii) state of completed processing denoted as 2.

In the event that each item was composed of 2 features, this categorization would fully represent the possible processing states.

Then the serial state space, $X_{sy} = \{(0, 0, 0), (0, 0, 1), (0, 0, 2), (0, 1, 0), (0, 1, 1), (0, 1, 2), (0, 2, 0), (0, 2, 1), (0, 2, 2), (1, 0, 0), (1, 0, 1), (1, 0, 2), (1, 1, 0), (1, 1, 1), (1, 1, 2), (1, 2, 0), (1, 2, 1), (1, 2, 2), (2, 0, 0), (2, 0, 1), (2, 0, 2), (2, 1, 0), (2, 1, 1), (2, 1, 2), (2, 2, 0), (2, 2, 1), (2, 2, 2)\}$.

Now the set of admissible states for a parallel system operating under these conditions is equal to the full range expressed above, that is $X_p = X_s$. Naturally, certain transitions are not permitted to occur. Thus, $(0, 0, 1) \rightarrow (0, 0, 2)$ or $(0, 0, 1) \rightarrow (1, 0, 1)$ may occur but not $(0, 0, 1) \rightarrow (0, 0, 0)$ or $(0, 0, 1) \rightarrow (0, 2, 1)$. The serial system is more constrained as its state space is $X_s = \{(0, 0, 0), (0, 0, 1), (0, 0, 2), (0, 1, 0), (0, 2, 0), (0, 2, 1), (0, 2, 2), (1, 0, 0), (1, 0, 2), (1, 2, 0), (1, 2, 2), (2, 0, 0), (2, 0, 1), (2, 0, 2), (2, 1, 0), (2, 1, 2), (2, 2, 0), (2, 2, 1), (0, 1, 2), (2, 2, 2)\}$.

Observe that states such as $(1, 1, 0)$ are disallowed (there are 7). Obviously the same remark as before applies to transition possibilities; something like $(0, 2, 2) \rightarrow (0, 1, 2)$ cannot occur.

Given these admissible states in a specific instance of a model, distinguishing a parallel processor from a serial processor becomes a matter of observing states that may be contained in either the serial or parallel state space set, but not in the intersection of the two. There are a number of other distinguishing systems-property aspects of parallel and serial processors as discussed by Townsend and Ashby (in press). Several of these aspects, though, simply are not observable under most practical instances or experiments we might consider. One defining property, though, which we shall make use of for actually testing purposes, and is noticeable from the example above, is that a serial processor may have only one item under partial completion for any specific time during the experiment. Obviously, if the processing of every item is definable by only two states, as when an item is comprised of only two features, then the state spaces will be identical for parallel and serial systems. With these facts in hand, we continue on to the next section where these properties are applied to actual test circumstances.

4. Methods and testing using second guessing

The systemic differences between parallel and serial processing when embedded in appropriate models in certain instances permit the design of experiments capable of testing parallel versus serial processing. As stated in Section 3, the most apparent discrepancies between the two types of processing systems are, in this case, the sets of admissible state spaces. (Put differently, certain *events* have probability 0 of occurrence in the sample space of serial, but not parallel models.)

For the purpose of actually observing subsets of these admissible state spaces it appears that ordinary reaction time experiments will simply not serve our purpose. The reason is that at most, item completion times are observed and the underlying intermediate states are hidden from view. Previous parallel-serial equivalence theorems are based on completion times, that is, are preferable to reaction time experiments (e.g., Townsend, 1972, 1976a, b; Vorberg, 1977). Therefore, we suggest a procedure involving accuracy with second guessing.

As an example here, we present a whole-report experiment with second guessing where a single display consists of two randomly chosen (with replacement) stimuli. This procedure was actually implemented and will be compared against the two tests developed later in this section. In these procedures we also restrict the second guess to be different from the first for each position in the display.

If we let C denote a correct guess and I denote an incorrect guess, all possible outputs of the experiment may be categorized by the four-dimensional vector consisting of:

(1st guess left position, 1st guess right position, 2nd guess left position, 2nd guess right position)

Given the binary output for each position, the admissible output set for the system is given by:

$$Y = \{(I, I, I, I), (I, I, I, C), (I, I, C, I), (I, I, C, C), (I, C, I, I), (I, C, C, I), (C, I, I, I), (C, I, I, C), (C, C, I, I)\}.$$

If we now denote the possible states for each position as in the last section where:

zero information state $\equiv 0$
 partial information state $\equiv 1$, and
 complete information state $\equiv 2$,

we now have the corresponding readout map between admissible states and possible (i.e., those with nonzero probability) outputs shown in Tab. 4A.

Now if we observe certain outputs a specific percentage of the time over many trials we are statistically 'assured' that certain states are actually occurring. As an example, we note that the output (I, I, C, C) may occur as a result of four distinct states. However, three of these states can only result in the output (I, I, C, C) by means of pure guessing. Given the stimulus set, we may determine a statistical maximum for the output (I, I, C, C) occurring as a result of the three pure guessing states. If the number of occurrences of this output is greater than this statistical maximum, we may be statistically confident of having observed the state denoted from Tab. 4A as (1, 1).

An even more general, but less apparent property in Tab. 4A is that the serial processor is permitted only one partially completed item. This property follows from the definition for a serial system and will be more thoroughly discussed in the application of Test 1 which we now present.

Tab. 4A: Example of state-to-output mappings for whole-report experiment with second guessing where the stimulus input consists of two items (left, right)

State map for two items to process (left, right) where		Readout map	possible outputs*
0 denotes no processing			
1 denotes partially processed state			
2 denotes totally processed state			
State number	(left, right)	States	
1	(0, 0)	<1, 2, 3, 4, 5, 6, 7, 8, 9>	→ (C, C, I, I)
2	(0, 1)	<1, 2, 4, 5, 7, 8>	→ (C, I, I, C)
3	(0, 2)	<1, 2, 4, 5, 7, 8>	→ (C, I, I, I)
4	(1, 0)	<1, 2, 3, 4, 5, 6>	→ (I, C, C, I)
5	(1, 1)	<1, 2, 3, 4, 5, 6>	→ (I, C, I, I)
6	(1, 2)	<1, 2, 4, 5>	→ (I, I, C, C)
7	(2, 0)	<1, 2, 4, 5>	→ (I, I, C, I)
8	(2, 1)	<1, 2, 4, 5>	→ (I, I, I, C)
9	(2, 2)	<1, 2, 4, 5>	→ (I, I, I, I)

* where C denotes correct guess
I denotes incorrect guess
and the vector of length four is (first response left position, first response right, 2nd guess left, 2nd guess right)

Test 1: Loose sieve: liberal nonparametric test of serial model employing second guess procedure

The null hypothesis employed for this test, as well as Test 2, is
 H_0 The system is a serial processor.

The alternative hypothesis is:

H_1 The system is not a serial processor.

The method used to test this hypothesis will be to statistically observe (as discussed previously) if any states occur which are not allowable states in a serial processor.

Using a whole-report, second-guess procedure as previously described, the parameters of such a procedure are as follows:

M = number of stimuli in stimulus set,

N = number of stimuli used in one display,

n = number of trials,

p = $1/(M-1)$ = probability of guessing correctly in any position on second guess given first response is incorrect,

J = number of trials on which the whole display is reported correctly on the first response,

K = number of items reported correctly on the first response in a display with N stimuli over n trials,

W = total number of correct guesses in the experiment.

The basic idea of this test is to statistically estimate the maximum number of correct second guesses possible in a serial model given the experimental paradigm described previously. We shall denote this maximum as $Y_{1-\alpha}$ (where $1 - \alpha$ is the probability of a confidence interval, explained further below) and compare it to the observed number of correct second guesses, W.

Within a whole report, second guess experimental framework, we have assumed that a completely processed item is identified correctly. The states that may yield an incorrect first response and a correct second guess on a particular item in a display are:

ment with second guessing where the

p

possible outputs*

- 6, 7, 8, 9) → (C, C, I, I)
- 8) → (C, I, I, C)
- 8) → (C, I, I, I)
- 6) → (I, C, C, I)
- 6) → (I, C, I, I)
- (I, I, C, C)
- (I, I, C, I)
- (I, I, I, C)
- (I, I, I, I)

notes correct guess

correct guess
tor of length four is (first
position, first response right,
t, 2nd guess right)

Model employing second guess

2, is

istically observe (as discussed
states in a serial processor.
eviously described, the para-

second guess given first response is

y on the first response,
isplay with N stimuli over n trials,

maximum number of correct
perimental paradigm described
ere $1 - \alpha$ is the probability
pare it to the observed number

ework, we have assumed that
es that may yield an incorrect
m in a display are:

- Type (1) — state where an item is as yet not even partially processed
- Type (2) — state where the item is partially processed.

We shall denote the maximum number of correct second guesses that result from type (1) state items as $Y_{1-\alpha}$ and the maximum number of correct second guesses that result from type (2) state items as Y'' . Therefore,

$$Y_{1-\alpha} = Y'_{1-\alpha} + Y''$$

Likewise, we shall let W' and W'' denote the number of second guesses in the experiment that occur from type (1) state items and type (2) state items, respectively. Therefore it follows that

$$W = W' + W''$$

With these equations, we have proposed that we will estimate the maximum number of second guesses possible from the additive combination of the maximum number of second guesses that may occur for items in the type (1) state and items in the type (2) state. Now, the number of items in the experiment which are incorrectly identified on the first response may be observed, and therefore considered a constant in the experiment. This total is some additive combination of the two disjoint types of states (i.e., type (1) and type (2) states). Our objective is to estimate the maximum number of second guesses that can occur, and it is more likely that an item in the state noted as type (2) will be correctly second-guessed. Therefore, we will first estimate the maximum number of type (2) states that can occur. We will then find the number of type (1) states that can occur by subtracting the maximum number of type (2) states that we have estimated from the constant number of total possibilities we have observed in the actual experiment.

With respect to a serial processor, the following property provides us with a constraint on the number of partially processed items an experiment may yield. We reiterate the

Serial systems axiom

Given a serial processor working on a single display of stimuli or items, the state of the system is such that at any time only one item may be in a state of partial completion.

Therefore, if there are n trials in an experiment, and on J trials the entire stimulus display is identified correctly, there remain a maximum of $(n - J)$ items in the type (2) state which may lead to a correct second guess. Now we require two assumptions regarding the item's second guessing probability. They are:

- (1) Pure guessing (i.e., type (1) state item) occurs when there is no information on an item. The structure of the guessing is such that independence exists across items in a display and across displays within an experiment.
- (2) For this test it is liberally assumed that any item partially processed (i.e., type (2) state items) but not correctly reported on the first response, is correctly guessed on the second response.

Now consider items in each display of type (1) where no processing has occurred. If no processing has been done, assumption 1 applies to the particular item. The independence assumption tells us we have a Bernoulli trial with the probability of correct second guessing being the parameter 'p'. Over n trials with N items shown on each trial the distribution on guessing with no information is binomial. To estimate the maximum

number of correct second guesses on these items where there is no information, we compute a statistical maximum for a binomially distributed variable. This distribution is characterized by two parameters, the total number of possible occurrences and the probability of an occurrence. In this instance an occurrence signifies a type (1) item which is not reported correctly on the first response. *The number of occurrences of type (1) items not reported correctly on the first response is*

$$(Nn - K - (n - J)),$$

where Nn is the total number of items presented during the experiment, K as denoted before is the number of items reported correctly on the first response, and $(n - J)$ is the maximum number of partially processed items not reported correctly on the first response in a serial processor. The probability of an occurrence or correct second guess in this instance is $p = 1/(M - 1)$ as denoted earlier.

If the possible number of occurrences is large, we may use a Normal approximation to estimate the maximum in a binomially distributed variable. In this instance we derive the statistical maximum from the positive side of the confidence interval

$$\bar{x} \pm z_{1-\alpha} \hat{\sigma},$$

where \bar{x} is the sample mean, $z_{1-\alpha}$ comes from the Normal table, and $\hat{\sigma}$ is the standard deviation. In our case, this yields a statistical maximum for the number of correct second guesses on type (1) state items which we have denoted $Y'_{1-\alpha}$ where

$$Y'_{1-\alpha} = p(Nn - K - (n - J)) + z_{1-\alpha} \sqrt{(Nn - K - (n - J)) p(1 - p)} \quad (4A)$$

and the probability of $(W' > Y'_{1-\alpha}) = \alpha$. (Recall that W' is the experimentally realized — but unobservable — number of type (1) second guesses.) In the above equation $p(Nn - K - (n - J))$ is the average number of items correctly second guessed estimated from the set of type (1) items which were not reported correctly on the first response, and $\sqrt{(Nn - K - (n - J)) p(1 - p)}$ is the standard deviation for these specific items.

Next, we must consider the second guessing capabilities for items in the type (2) state; the partially processed state. Remember that the serial systems axiom permits at most one partially processed item per display or trial for each display not correctly reported on the first response, and that assumption (2) allows all of the partially processed items to be correctly guessed on the second response. We can then calculate the maximum number of correct second guesses for partially processed items as being $Y'' = (n - J)$. Note that the probability of $(W'' < Y'') = 1$ due to our assumptions about partially processed items.

Now we combine the estimate for maximum second guessing on items where there is partial information, $Y'' = (n - J)$, with our Normal approximation on items where there is no information, Y' . Our results yield

$$Y_{1-\alpha} = Y'_{1-\alpha} + Y'' = p(Nn - K - (n - J)) + (n - J) + z_{1-\alpha} \sqrt{(Nn - K - (n - J)) p(1 - p)},$$

and that $P(W = W' + W'' \leq Y_{1-\alpha} = Y'_{1-\alpha} + Y'') = 1 - \alpha$.

Equivalently, with confidence $1 - \alpha$, the maximum number of correct second guesses in the experiment will be less than $Y_{1-\alpha}$.

Test 2: Tighter sieve: conservative nonparametric test employing second guess procedure

Test 2 is a refinement of Test 1. In Test 1 it was assumed that partial information always led to a correct second guess if the first guess was incorrect. Under this initial assumption we found that the type (2) term of the estimate for the maximum number of second guesses was $Y'' = (n - J)$. This term may tend to yield gross overestimates of second guessing ability, so we now wish to make stronger assumptions about the amount of partial information a Subject has and the usefulness of the information.

We begin with a restriction on the stimulus set. Assume a stimulus set composed of the power set of Q feature positions or element positions. This stimulus contains 2^Q stimuli, including a blank stimulus. We now suppose that the Subject identifies a stimulus by searching or sampling from each possible feature position. Therefore the Subject must check Q feature positions in order to positively identify a stimulus.

Our next assumption is that given partial processing, we assume it equally likely that the Subject has processed anywhere from one to $Q - 1$ of the feature positions. Therefore, the average number of feature positions processed is given by:

$$\left(\frac{1}{Q-1}\right) \sum_{k=1}^{Q-1} k = \left(\frac{1}{Q-1}\right) \left(\frac{(Q-1)Q}{2}\right) = \frac{Q}{2}.$$

Using the term, $Q/2$, as the average number of feature positions processed given partial processing, we wish to estimate the probability of making a correct second guess given an incorrect first response on an item when $Q/2$ feature positions have been processed. For a stimulus set made up from a power set of Q feature positions, an optimal decision process is one that reduces the response set by a factor of two for each feature position processed. The reader may recall that in a stimulus set made up of the power set of features, half of the stimuli contain a specific feature and half of them do not. Therefore, when a feature position is processed, half of the stimulus set may be rejected for response purposes on the grounds that they do or do not contain a feature in a specific position.

Thus, given a stimulus set composed of the power set of Q features where one feature position is processed and the first response on a single item is incorrect, there remain $[2^{(Q-1)} - 1]$ possible items in the stimulus set from which to choose. Since the average number of features processed in our case is $Q/2$, there remain $[2^{(Q/2)} - 1]$ stimuli remaining in the response set and the probability of guessing correctly from this set is $1/(2^{(Q/2)} - 1)$.

Now replacing the type (2) term of Test 1, Y'' , with the new guessing factor for partial information and multiplying this factor by the appropriate number of trials where a partially processed item is available for second guessing, we obtain the new estimate for the $(1 - \alpha)$ confidence statistic on the maximum number of second guesses where:

$$y_{1-\alpha} = y'_{1-\alpha} + y'' = p(Nn - K - (n - J)) + \frac{(n - J)}{(2^{Q/2} - 1)} + z_{1-\alpha} \sqrt{(Nn - K - (n - J)) p(1 - p)},$$

and now

$$P(\bar{w} \leq y_{1-\alpha}) = 1 - \alpha,$$

or equivalently, with confidence $1 - \alpha$, the maximum number of correct second guesses in the experiment will be less than $y_{1-\alpha}$.

Whole report experiment utilizing Test 1 and Test 2

Townsend and Evans conducted an accuracy experiment of the whole report variety employing the second guess procedure (see Townsend, 1981 for further discussion of issues in whole report behavior). The stimulus display for each trial consisted of two items displayed on a cathode-ray tube computer terminal with lateral and backward masking to reduce visibility. Three subjects reported first guesses and second guesses which were restricted to being different from the first guesses in each of the two display positions. The stimuli for the two positions in each display were randomly chosen *with replacement* from the alphabet minus the letters {A, E, I, O, U, Y}. For the purpose

Tab. 4B: TOWNSEND and EVANS whole-report results¹

N = number of letters in a stimulus display = 2

M = number of letters in the stimulus set = 19

P = probability of a correct second guess given the first response is incorrect and no information is known about the stimulus = $1/(M - 1) = 1/18$

Statistic or parameter	subject		
	1	2	3
n	1600	1152	1728
J	140	249	473
K	904	995	1714
(Nn - K - (n - J))	836	406	487
n - J	1460	903	1255
$z_{.95} \sqrt{(Nn - K - (n - J)) p(1 - p)}$	59	32	37
Total number of observed correct 2nd guesses	199	130	268
Test 1 (maximum statistic)**	1820	935	1292
Test 2 (maximum statistic)** with Q			
= estimated number of feature positions =	4	547	333
	5	373	226
	6	269	161
	7	201	120*

* rejects serial model for this level

** maximum statistical estimate at 95 per cent confidence level (i.e., using Z (alpha = .05))

¹ Note that the numbers listed under Tests 1 or 2 were compared with observed total number of correct guesses. The fact that the latter quantity is usually less than the test criteria indicates the general passing of the tests.

of brevity we shall dispense with the rest of the procedural description. The obtained statistics in reference to Test 1 and Test 2 are given in Tab. 4B.

Since the stimulus set used was not actually a power set of feature positions, Q, as taken from Test 2, was made a free parameter. Tab. 4B shows that the serial model was not rejected for any of the subjects using Test 1. Under Test 2, the serial model was rejected for Subject 3 when Q, the number of feature positions, was set to six or above, and for Subject 2 when Q was set to seven or above.

With only two positions to be processed, it might have been expected that parallel processing, or at least some alternative to serial processing might have taken place. However, we might note that Test 1 is ultra-liberal with respect to accepting a serial processor when only two stimulus positions exist in a display. Test 2 may also be con-

ment of the whole report variety end, 1981 for further discussion play for each trial consisted of two terminal with lateral and backward and first guesses and second guesses guesses in each of the two display display were randomly chosen with {E, I, O, U, Y}. For the purpose

onse is incorrect and no information is

ect

2	3
1152	1728
249	473
995	1714
406	487
903	1255
32	37
130	268
935	1292
333	456
226	307
161	217*
120*	

el (i.e., using Z (alpha = .05))

ed with observed total number of correct test criteria indicates the general passing

cedural description. The obtained Tab. 4B.

er set of feature positions, Q, as 4B shows that the serial model . Under Test 2, the serial model ature positions, was set to six or ove.

have been expected that parallel processing might have taken place. with respect to accepting a serial display. Test 2 may also be con-

sidered somewhat liberal (albeit much less so than Test 1) due to our assumption that the subjects operate in an optimal fashion with respect to guessing with partial information. The authors feel that the algorithm used by the <IPS> to arrive at an output is of importance in this context. In these two tests, we have chosen two separate algorithms,

(1) the <IPS> can respond correctly on the second guess anytime it has any information at all about the item, and

(2) the <IPS> works optimally with respect to a discrete power set of feature position information.

The stimulus set used in the experiment presented was not particularly designed for the study of (2), which we feel is an attractive option open for study. Further research into (2) as a viable decision or response algorithm with respect to a whole-report, second guess paradigm can be conducted by choosing various stimulus sets composed of power sets of various feature positions.

In summing up, we must conclude that the data have successfully navigated the first two nonparametric serial sieves. This suggests the existence of a detailed serial model that might predict the finer grain of the experimental results; one, for instance, based on a specific stochastic process.

It should be noted that we have not ruled out parallel processing with the current sieves; we simply have not been able to reject serial processing. However, it should be noted that with regard to the present type of experiment, the serial models make much stronger assumptions than parallel models. In fact, it has been suggested that many hybrid models possess an underlying event space like those of parallel models; but, few possess the tightly restricted event space of serial models (Townsend & Ashby, in press). Nevertheless, it should be possible to construct nonparametric sieve tests for parallel processing analogous to those constructed for serial models.

References

- ANDERSON, J. R.: Language, memory, and thought. Hillsdale, N.J.: Lawrence Erlbaum Assoc., 1976
- ARBIB, M. A.: A common framework for automata theory and control theory. *J. SIAM Control*, 1965, 3, 207-222
- ATKINSON, R. C., HOLMGREN, J. E., & JUOLA, J. F.: Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics*, 1969, 6, 321-326
- BERTALANFFY, L. VON: *Theoretische Biologie*, vols. 1 & 2. Berlin: Borntraeger, 1932, 1942; 2nd ed., Bern: Francke, 1951
- BOOTH, TAYLOR L.: *Sequential machines and automata theory*. New York: Wiley, 1967
- BROADBENT, D. E.: World frequency effect and response bias. *Psychological Review*, 1967, 74, 1, 1-15
- CATTELL, J. MCK.: The time taken up by the cerebral operations. *Mind*, 1886, 11, 220-242, 377-392, 524-538
- DiSTEFANO, J. J., STUBBERUD, A. R., & WILLIAMS, I. J.: *SCHAUMS's outline of theory and problems of feedback and control systems*. New York: Wiley, 1967
- DONDERS, F. C.: On the speed of mental processes. In: W. G. KOSTER (Ed. and trans.), *Attention and Performance 2*. Amsterdam: North-Holland, 1969
- EGETH, H.: Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics*, 1966, 1, 242-252
- ERIKSEN, C. W., & SCHULTZ, D. W.: Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, 1979, 25, 249-263
- ESTES, W., & TAYLOR, A.: A detection method and probabilistic model for assessing information processing from brief visual displays. *Proceedings of the National Academy of Science*, 1964, 52, 446 to 454

- FÄLMAGNE, J. C., & THEIOS, J.: On attention and memory in reaction time experiments. *Acta Psychologica*, Amsterdam 1969, 30, 316—323
- GREENO, J. G., & STEINER, T. E.: Markovian processes with identifiable states: General considerations and applications to all-or-none learning. *Psychometrika*, 1964, 29, 4, 309—333
- HICK, W. E.: On the rate of gain of information. *Quarterly J. of Experimental Psychology*, 4, 1952, 11—26
- HOFFMAN, JAMES E.: Search through a sequentially presented visual display. *Perception & Psychophysics*, 1978, 23, 1, 1—11
- HYLAN, J. P.: The distribution of attention. *Psychological Review*, 1903, 10, 493—533
- KALMAN, R. E.: Algebraic structure of linear dynamical systems. I. The module of Σ . *Proceedings of the National Academy of Science (USA)*, 1965, 54, 1503—1508
- KALMAN, R. E., ARBIB, M. A., & FALB, P. L.: *Topics in Mathematical System Theory*. New York: McGraw-Hill, 1969
- KANTOWITZ, B. H.: Double stimulation. In: B. H. KANTOWITZ (Ed.), *Human Information Processing: Tutorials in Performance and Cognition*, Potomac, Md.: Lawrence Erlbaum Associates, 1974
- KLIR, GEORGE J. (Ed.): *Trends in general systems theory*. New York: Wiley & Sons, 1972
- MCCLELLAND, JAMES L.: On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 1979, 86, 4, 287—330
- MCGILL, W. J., & GIBBON, J.: The general-gamma distribution and reaction times. *J. of Mathematical Psychology*, 1965, 2, 1—18
- MESAROVIC, M. D.: Systems theory and biology-view of a theoretician. In: M. D. MESAROVIC (Ed.), *Systems Theory and Biology*. New York: Springer-Verlag, 1968
- NEISSER, U., NOVICK, R., & LAZAR, R.: Searching for ten targets simultaneously. *Perceptual and Motor Skills*, 1963, 17, 955—961
- PACHELLA, ROBERT G.: The interpretation of reaction time in information-processing research. In: B. H. KANTOWITZ (Ed.), *Human Information Processing: Tutorials in Performance and Cognition*, Potomac, Md.: Lawrence Erlbaum Associates, 1974
- PACUT, A.: Mathematical modelling of reaction latency. Technical Report, Division of Applied Mathematics, Brown University, 1981
- SNODGRASS, J. G., & TOWNSEND, J. T.: Comparing parallel and serial models: Theory and implementation. *J. of Experimental Psychology: Human Perception and Performance*, 1980, 6, 2, 330—354
- STERNBERG, S.: High-speed scanning in human memory. *Science*, 1966, 153, 652—664
- STERNBERG, S.: Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics*, 1967, 2, 45—53
- STERNBERG, S.: The discovery of processing stages: Extensions of Donders' method. In: W. G. KOSTER (Ed. and trans.), *Attention and Performance 2*. Amsterdam: North-Holland, 1969
- TAYLOR, D. A.: Stage analysis of reaction time. *Psychological Bulletin*, 1976, 83, 2, 161—191
- THEIOS, J.: Reaction time measurements in the study of memory processes: Theory and data. In: G. H. BOWER (Ed.), *The Psychology of Learning and Motivation*. New York: Academic Press, 1973
- TOWNSEND, J. T.: A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, 1971, 10, 161—163
- TOWNSEND, J. T.: Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical & Statistical Psychology*, 1972, 25, 168—199
- TOWNSEND, J. T.: 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*, Potomac, Md.: Lawrence Erlbaum Associates, 1974
- TOWNSEND, J. T., & SNODGRASS, J. G.: A serial vs. parallel testing paradigm when "same" and "different" rates differ. Paper presented to Psychonomic Society, Boston, 1974
- TOWNSEND, J. T.: A stochastic theory of matching processes. *J. of Mathematical Psychology*, 1976a, 14, 1—52
- TOWNSEND, J. T.: Serial and within-stage independent parallel model equivalence on the minimum completion time. *J. of Mathematical Psychology*, 1976b, 114, 3, 219—238
- TOWNSEND, J. T.: Some characteristics of visual whole report behavior. *Acta Psychologica*, 47, 2, 1981, 149—173
- TOWNSEND, J. T., & ASHBY, F. G.: *Stochastic modeling of elementary cognitive processes*, Cambridge University Press, in press

ion time experiments. *Acta Psychologica*, 1973, 10, 493—533

ifiable states: General considerations. *Journal of Experimental Psychology*, 1952, 45, 29, 4, 309—333

of Experimental Psychology, 4, 1952, 45, 29, 4, 309—333

visual display. *Perception & Psychophysics*, 1973, 10, 493—533

I. The module of Σ . *Proceedings of the International System Theory Conference*, 1973, 10, 493—533

d.), *Human Information Processing: A Practical Approach*. Erlbaum Associates, 1974

rk: Wiley & Sons, 1972

s: An examination of systems of processes. *Journal of Experimental Psychology*, 1973, 10, 493—533

d reaction times. *Journal of Mathematical Psychology*, 1973, 10, 493—533

etician. In: M. D. MESAROVIC (Ed.), *Mathematical Psychology*. Wiley, 1973, 10, 493—533

multaneously. *Perceptual and Motor Skills*, 1973, 10, 493—533

formation-processing research. In: *Information Processing in Man-Machine Systems*. Academic Press, 1973, 10, 493—533

Report, Division of Applied Mathematics, 1973, 10, 493—533

erial models: Theory and implementation. *Journal of Experimental Psychology*, 1980, 6, 2, 330—354

1966, 153, 652—664

vidence from reaction-time measurements. *Journal of Experimental Psychology*, 1966, 153, 652—664

Donders' method. In: W. G. KOSTER (Ed.), *Psychological Processes*. North-Holland, 1969

ulletin, 1976, 83, 2, 161—191

rocesses: Theory and data. In: G. H. Bower (Ed.), *Information Processing in Man-Machine Systems*. New York: Academic Press, 1973

rial processes. *Perception & Psychophysics*, 1973, 10, 493—533

parallel and serial processes. *British Journal of Psychology*, 1973, 10, 493—533

a finite number of inputs. In: B. H. Welford (Ed.), *Information Processing in Man-Machine Systems*. Potomac, 1973, 10, 493—533

g paradigm when "same" and "different". *Journal of Experimental Psychology*, 1974, 10, 493—533

of Mathematical Psychology, 1976a, 10, 493—533

el equivalence on the minimum comparison. *Journal of Experimental Psychology*, 1976a, 10, 493—533

ivor. *Acta Psychologica*, 1976a, 10, 493—533

tary cognitive processes, Cambridge

TOWNSEND, J. T., & LANDON, D. E.: Mathematical models of recognition and confusion in psychology. *International J. of Mathematical Social Sciences* (in press a)

TOWNSEND, J. T., & LANDON, D. E.: An experimental and theoretical investigation of the constant-ratio rule and other models of visual letter confusion. *J. of Mathematical Psychology* (in press b)

VORBERG, D.: On the equivalence of parallel and serial processes. Paper presented at Mathematical Psychology Meetings, Los Angeles, 1977

WEAVER, W.: Science and complexity. *American Scientist*, 1948, 36, 536—544