




# A beginning quantitative taxonomy of cognitive activation systems and application to continuous flow processes

James T. Townsend<sup>1</sup> · Michael J. Wenger<sup>2</sup> 

Accepted: 12 October 2020 / Published online: 7 January 2021  
© The Psychonomic Society, Inc. 2021

## Abstract

Much progress has been made in the investigation of perceptual, cognitive, and action mechanisms under the assumption that when one subprocess precedes another, the first one starts and finishes before the other begins. We call such processes “Dondersian” after the Dutch physiologist who first formulated this concept. Serial systems obey this precept (e.g., Townsend, 1974). However, most dynamic systems in nature do not: instead, each subprocess communicates its state to its immediate successors continuously. Although the mathematics for physical systems has received extensive treatment over the last three centuries, applications to human cognition have been exiguous. Therefore, the pioneering papers by Charles Eriksen and colleagues on continuous flow dynamics (e.g., Eriksen & Schulz, *Perception & Psychophysics*, 25, 249–263, 1979; Coles et al., *Journal of Experimental Psychology: Human Perception and Performance*, 11(5), 529, 1985) must be viewed as truly revolutionary. Surprisingly, there has been almost no advancement on this front since. With the goal of bringing this theme back into the scientific consciousness and extending and deepening our understanding of such systems, we develop a taxonomy that emphasizes the fundamental characteristics of continuous flow dynamics. Subsequently, we complexify the treated systems in such a way as to illustrate the popular cascade model (Ashby, *Psychological Review*, 89, 599–607, 1982; McClelland, *Psychological Review*, 86, 287–330, 1979) and use it to simulate the classic findings of Eriksen and colleagues (Eriksen & Hoffman, *Perception & Psychophysics*, 12(2), 201–204, 1972).

**Keywords** Math modeling · Perception and action · Reaction time methods

Donders 1868/1969 and many other early scientific psychologists conceived of the mind as a series of more or less independent subsystems, with nonoverlapping, discrete processing durations or stages (Sanders, 1977; Sternberg, 1969; Townsend, 1972; Van Zandt & Townsend, 2012). The non-overlapping stages condition was typically complemented through Donders’ axiom of pure insertion, wherein it is assumed that a psychological process can

be inserted or withdrawn through manipulation of the psychological task (Ashby & Townsend, 1980; Roberts & Sternberg, 1993; Townsend & Ashby, 1983; Zhang, Walsh, & Anderson, 2018). Such an idea is closer to the activities of a middle 20th-century digital computer’s accumulator operations than a series of simple electrical components with lumped systems properties; or the apparent neural structures and dynamics linking cortical regions; or, indeed, than the analogue connectionist and distributed memory of modern neuralistic theorizing (Anderson, 1995; Ashby, 1952, 1982; Grossberg, 1988; Hertz et al., 1991; McClelland 1979). We denote this type of architecture and information processing a continuous flow system.

Discrete flow systems have been the overwhelming choice in human information processing and cognition. The history of psychology, reaching back into the early days of philosophy, has always counterposed more analytic, atomistic conceptions of perception, cognition and action with more global, intuitive, flowing and dynamic approaches. One can find these opposing forces depicted in any history of scientific psychology (e.g., Boring, 1957). Perhaps

---

An early version of this theory was originally published as: Townsend, J. T. & Fikes, T. (1995). A beginning quantitative taxonomy of cognitive activation systems and application to continuous flow processes. Cognitive Science Research Report 131, Indiana University. This paper would not have been possible without the encouragement and support throughout the project by Action Editor Joseph Lappin.

---

✉ James T. Townsend  
jtownsen@indiana.edu

<sup>1</sup> Indiana University, Bloomington, IN, USA

<sup>2</sup> The University of Oklahoma, Norman, OK, USA

because of their relative ease of formal analysis, the discrete and finite seem to attract mathematical theory and sometimes, even empirical assessment at an earlier and greater rate than the more global, infinite, and intuitive. The latter, even in the present day, sometimes seem to languish more in the world of verbal description and metaphor.

This propensity in behavioral science has persevered despite the ubiquity of continuous flow systems in other sciences and engineering. We shall cite other investigators who employed continuous flow concepts in their narrative below but first we call out the pioneer who not only foresaw the necessity of exploring the latter types of systems, but even offered experimental evidence to support the concept of continuous flow. That pioneer was, of course, Charles W. Eriksen. He and his colleagues, using the fairly primitive tools of electromyography, showed that the nature of surrounding flankers influenced the human motor response throughout its trajectory (Coles et al., 1985; Eriksen & Hoffman, 1972, 1973, 1974, 1979).

It should be a given that some important brain processes involve continuous flow mechanisms. For instance, dynamic systems have a history of utility in development, motor control, and neuroscience (see, discussions in, e.g., Haykin & Fuster 2014; Thelen & Smith, 1994). Given the certainty that at least some tasks and particular subsystems of the information processing chain obey continuous rather than discrete flow, it may seem odd, outside of the above historical and epistemological comments, that so little theoretical and even empirical work has been accomplished. Further, the continuous flow approach might have been thought to be the first choice even of an early cognitive theorist in the nineteenth century, due to their visibility in the hard sciences and engineering of the period, and the fact that many embryonic psychologists of the nineteenth century were physicists or physiologists. The continuous flow notion is also much more consonant with the more recent dynamic systems initiative in cognitive science (Bingham, Schmidt, & Rosenblum, 1995; Busemeyer & Townsend, 1993; Heath & Fulham, 1988; Molenaar, De Gooijer, & Schmitz, 1992; Port & van Gelder, 1995; Thelen & Smith, 1994; Patterson et al., 2013).

Rigorous mathematical treatment of the construct of continuous flow in psychology is rare, but there are nonetheless several key studies. A study notable both for its innovation as well as deriving some results concerning selective influence predictions on response time was published by Taylor (1976). That paper considered the possibility that successive sub-processes could overlap in their processing times.

Perhaps the first paper to commit to a continuous flow across stages was McClelland's cascade model (McClelland, 1979) embellished with the addition of a Gaussian random variable acting as decision threshold. The cascade model was, despite belonging to a classical continuous flow

type of dynamic system, shown to be capable of making approximately additive predictions usually associated with Dondersian structures. The original Cascade Model was followed up by a commentary and allied derivations from F.G. Ashby 1982. The cascade model represents neither the simplest and most constrained, nor the most general, continuous flow type of system. However, it does exemplify several of the important non-trivial characteristics that a dynamic system might possess.

In what follows, we consider these characteristics and provide a taxonomy of system aspects that we hope will prove to be valuable for the study of cognition, motor control, perception, memory search, and categorization and decision making. The style of our presentation is inclined toward the systematic as opposed to the colloquial. Our own experience with quite complex material suggests to us that, with so many technical distinctions, a too-colloquial fashion can make it very challenging when engaging with the material for comprehension, cross-checking and research purposes. We try to ameliorate this parched rigor with some examples and toward the end, an application and look to the future.

## Characteristics of activation systems

### Structures and skeletons

By structure, we mean a graph that shows which hypothesized subsystems (= subprocesses) connect with each other and the order of processing among those pictured. It is critical to keep structure separate from other notions such as the type of state space, the mode of information flow, and the like. Note that in this sense, serial systems and sequential, continuous flow systems satisfy the same type of structure but not the same architecture, because we reserve the term serial architecture to include the postulate of a Dondersian, discrete flow (see, e.g., Townsend & Ashby, 1983, chapter 2).

Our targeted structures will be the simplest one can think of. They form a simple so-called total order of the underlying subprocesses (also called a linear order or simple order). A total order assumes transitivity, connectedness and antisymmetry. In such a skeleton, assume  $p, q, r$  are subprocesses in the set  $A$ . Then transitivity implies that if  $p \geq q$  and  $q \geq r$ , then also  $p \geq r$ . Connectedness implies that if  $p, q$  are in  $A$  then either  $p \geq q$  or  $q \geq p$ . Antisymmetry is the condition that if both  $p \geq q$  and  $q \geq p$  then  $p = q$ . We point to a very large and natural set of structures where everything flows forward so that transitivity is in force, antisymmetry also occurs and reflexivity is true (i.e.,  $p \geq p$ ) but connectedness does not hold in general. The last means that some branches in the

structure are not ordered relative to one another. This is the notion of a partial order (that is, the relations of transitivity, reflexivity and antisymmetry hold). They will be a natural generalization of our simple forward flow systems.

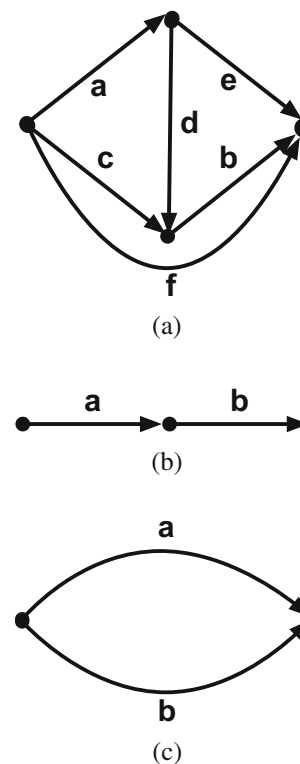
The appearance of such structures is of a set of subprocesses which flow from start to finish, and where individual processes may, but need not, have connecting paths between them. Parallel and serial systems are special cases, and no such network permits feedback to occur (Townsend & Ashby, 1983, chapter 1). In such systems, the engineer or scientist defines to what the final ‘stop’ node refers. In a psychological model, it might mean the actual pushing of a reaction time button, the point at which information accrual advances beyond a decisional threshold or when some other presumably punctate event occurs. The Eriksen and colleagues’ novel design and method of measuring the hand’s trajectory may auger a new dawn of strategies for systems-identification that go far beyond the traditional response times and accuracy measures. The fairly recent experimental procedures oriented toward analysis of continuous motor trajectories, represented in the present collection by Erb, Smith, and Moher (2020) are perhaps the most obvious and natural inheritors of the Eriksen continuous flow tradition.

When supplemented with assumptions about flow and stopping rules that specify when a subprocess can get underway, and characterization of subprocess durations, behavioral or physiological predictions can be made. For instance, can a subprocess begin when any of several paths ending on its start node are finished, only when all predecessors are complete or some other more complex rule), actual systems emerge. The PERT (program evaluation and review technique; an acronym drawn from operations research networks referred to above) are partially ordered structures which assume discrete time processing among constituent processes and typically an exhaustive stopping rule between the just-preceding processes and the subsequent processes into which they feed. They have been thoroughly investigated as models of cognitive processing times by Schweickert and colleagues (Fisher & Goldstein, 1983; Schweickert, 1978, 1989, 2012).

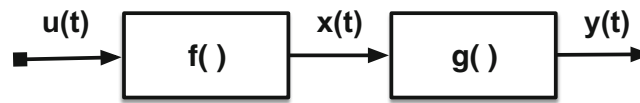
The individual processes may be deterministic or stochastic and the latter are, of course, more realistic. Figure 1a shows a prototypical PERT network called an embellished Wheatstone bridge (Schweickert, 1989; Townsend & Schweickert, 1989). This network is kind of a microcosm for the layout of subprocesses *a* and *b*, since there is a path through both *a* and *b*, a path that goes through either without passing through the other, and a path that includes neither. Typically, PERT networks have been endowed with an exhaustive stopping rule (all processes

feeding into a subsequent node and its process must be done before the succeeding one can begin and there is no overlap of processing times (Schweickert, 1978) but other stopping rules are readily available (Schweickert, Fisher, & Sung, 2012).

As noted above, serial (Fig. 1b) and parallel (Fig. 1c) systems are special cases of PERT networks. However, a sequentially arranged architecture does not imply seriality because, as observed above, the latter also demands that the processing times of the successive stages not overlap whereas a sequential architecture could have imposed on it, a continuous flow, or other hybrid type of processing (see below). The same is true of more complex networks, in that an investigator may use a PERT-like skeleton to represent the architecture, but assume a form of continuous, rather than discrete, flow processing. Our concentration in this paper is on a single sequential structure and flow of processes, as in Fig. 2. Feedback is allowed within each constituent sub-process but we do not allow feedback between them.



**Fig. 1** PERT network representations of three system architectures are shown. Under the assumptions of exhaustive processing and discrete time processing (i.e., that a process feeds into subsequent processes in an all or none fashion, at a single instant in time, and that subsequent processes may not begin processing until all preceding processes have finished). **a** An embellished Wheatstone bridge with concurrent and sequential processes, **b** two serial processes, **c** two parallel processes



**Fig. 2** A simple two-component, sequential architecture. In Table 1, three information processing models from the literature with this general architecture are summarized within our taxonomy. In our terminology,  $u$ ,  $x$ , and  $y$  are time-varying states:  $u$  represents the input to the system,  $x$  is the result of some transformation on this input by

the operator  $f()$ , and  $y$  is the result of some transformation on  $x$  by an operator  $g()$ . Boxes therefore represent operators, while arrows represent the transmission of state variables (which may or may not be associated with time lags)

**Characterization of system state space and input and output spaces**

A minimally complete description of a processing system will include the structure, the time variable, an input space, and a state space for the state of processing or activation associated with each subsystem. Typically, the output of a subsystem will be taken as the specification of the state of the subsystem at the juncture.

In the general purview here, we desire that the state and time descriptions be allowed to be different, although our specific taxonomy will be confined to continuous flow systems. We conceive of systems where the input and outputs, and possibly the states, are continuous, and yet the state transitions may be discrete. A good example is a Dondersian system where each subsystem acts as a separate temporal stage and overall processing is thus strictly serial with each successive stage only beginning when the previous one is completed (e.g., Townsend, 1974). We can mention this type of system in terms of our present discussion of forward-flow models. However, such constructions are outside most of our taxonomy because they require concepts like temporal lag to make them obey Dondersian principals. Therefore, after a little more discussion of these to aid in the comparison with the corpus

of continuous flow models, we shall drop them from our descriptions (but see, e.g., Townsend and Ashby, 1983, pp. 401-409).

Let us call this *temporally discrete transmission*. Certainly, we do not want to rule out the possibility that the underlying states are continuous even though processing is discrete in this sense (see, e.g., Table 1). As noted by J. Miller 1988, processing internal to the distinct subsystems could still lie on a continuum and the information transmitted (discretely in time) to the next stage, could be continuous too (e.g., any positive or negative number). When transitions are continuous among subsystems, the information flow is going on between two such subsystems all the time, at least on a given time interval. If a system is also *lumped*, then information is transmitted throughout the subsystems without any lag. The usual theory of ordinary differential equations does not include reference to systems with lag (see, e.g., Cunningham, 1958; MacDonald, 2013), although they abound in many physical (and certainly neural) phenomena. Lagged systems call for the use of difference-differential equations, differential-integral equations or other rather arcane devices (see the references above for certain techniques and applications to biology). We will witness a form of lag in Table 1, in an example of a Dondersian system where the stage states are

**Table 1** Summary of the taxonomy dimensions and example models

	Model		
	Dondersian	Composite	Cascade
Architecture	Townsend (1974) sequential	Schweickert (1989) sequential	McClelland (1979) sequential
Independent variable	time	time	space
Nature of state spaces	Discrete or continuous in state; continuous in time: $t, x, y \in \mathbb{R}^+$	Continuous in state and time: $t, x, y \in \mathbb{R}^+$	Continuous in state and time: $t, x, y \in \mathbb{R}^+$
System memory	Memoried	Memoryless	Memoried
Stochasticity	Deterministic	Deterministic	Deterministic
Example equation	$u = 0, t < 0$ $\frac{dx}{dt} = -a(x - u)$ $\frac{dy}{dt} = -U(x - c_x)b[y - x(t - t_x)]$ $t_x = \text{first time s.t. } t_x = c_x$ $t_y = \text{first time s.t. } Y(t_y) = c_y$	$x = f(u, t)$ $y = g(x)$	$\frac{dx}{dt} = f(x, u) = -a(x - u)$ $\frac{dy}{dt} = g(y, x) = -b(y - x)$

basically continuous, but transmission occurs only when the preceding state reaches a criterion. Otherwise, our systems will be of the lumped variety.

The class of Dondersian systems does not usually specify the details of processing, in particular, there is usually lacking a state space of activation. All we assume is that whatever they are, the subsequent processes are nonoverlapping in their operations. However, one subset of Dondersian systems could conceive of the state space as discrete with the “clumps” of information passed on in their entirety, within a general framework of discrete automata (e.g., Townsend & Ashby, 1983, pp. 409–412). However, another conception is that the state space is continuous but that a lag is introduced in transmission from one state to the next, that is greater than or equal to the processing duration of the present stage (see Table 1 and Townsend & Ashby, 1983, p. 408). Thus, delayed, but continuous, systems theory gives one approach to the study of a form of Dondersian systems theory that may be more realistic than discrete automata theory for some situations. Additional discussion of lag and other generalizations of these dimension will be visited in the Discussion. We now proceed to set up a rigorous basis for describing the dynamics of sequentially arranged systems. It follows closely the terminology of systems theory such as Klir (1969) and Booth (1967) and Padulo and Arbib 1974. It can readily be expanded to encompass more general architectures.

We first describe the critical parts that will make up our continuous flow dynamic systems. Our next step will be to assemble the systems more or less in order of increasing complexity. Patently, systems based on differential equations are more particular than those not making differentiability assumptions. It will prove propitious to specialize each of the cases to differential equations after the more general dynamics are considered.

Let the input space be  $U$  and the output space be designated  $Y$  with the time set being  $T$ . Let the state space of a first subsystem be  $X$  and that of a second subsystem,  $Y$ . That is,  $Y$  designates both the state of the second subsystem as well as the final output. Some situations might demand yet a third space for a distinct set of outputs determined by  $Y$ . We shall not need that extension. Lower case letters indicate specific values. The transition operator that leads from the input to a state in  $X$  is  $F$ , that for  $Y$  is  $G$ . Let us assume that in some cases, the state itself can help determine the next state (which occurs instantaneously in a continuous state system). In actual cases, this dependence is often seen in the local description given by a difference or differential equation. In most realizations of such systems,  $Y$  will be compared with a decision threshold that, when met or exceeded, drives a decision and response. However,  $Y$ , or its facsimile, could also drive a continuous response, for example, something like a hand-in-motion-toward a choice

between two responses. This is exactly what we'll see in our Eriksen flanker model later on.

The systems we consider here allow an initial condition on  $X$  but, not on  $Y$ . And so here we need to address a set of initial (no pun intended) questions. First, does a subsystem age, meaning does the subsystem change in its characteristics simply as a function of time? We refer to an ageless subsystem as *time invariant* otherwise, *time variable*. Second, is there an input  $u(t)$  to the  $X$  subsystem? In most cases, we will assume so. Third, is the evolution of a subsystem a function of its present state? The very simplest of systems may not be but as we will see, most will be. In fact, it will be shown that, in general such a property implies a type of memory in the subsystem.

We begin with the simplest systems, in which there is no input, and the next immediate state is a function only of input and/or time but not the current state. Thus, they are memoryless, and they are time invariant in both subprocesses. The resulting map is

$$F : T \rightarrow X$$

meaning that  $F$  is a direct function (only) of time  $t$  or equivalently,  $x = F(t)$ . But then, instantaneously, the  $Y$  subsystem maps the state  $x$  into a particular value  $y$ , that is,

$$G : X \rightarrow Y,$$

or in everyday notation,  $y = G(F(t))$ . This is the most elementary dynamic with which we deal and constitutes the class of systems investigated in the seminal work of Schweickert (Schweickert, 1989; Schweickert & Mounts, 1998). These are known as composite functions. We can trivially insert an input  $u(t)$  as in

$$F : T \times U \rightarrow X$$

and

$$G : X \rightarrow Y$$

In addition, it is a simple task to let such a subsystem also depend on an initial condition, which we implement here only for the  $x$  subsystem. The result is a *generalized compositional system*. Here,  $x = F(t, u, x_1)$  and  $y = G(t, x) = G[t, F(t, u, x_1)]$ . Now  $Y$  is an actual function of the single time  $t$ , and the exact values of  $u, x$  at that time, plus an initial condition specified by  $x_1$ . Certainly,  $G$  is now an ordinary function of the three variables  $t, x$  and  $y$ . In the special case that  $G$  is a function only of  $x$  over the time interval but not separately of  $t$ , then we obtain  $y = G[F(t, u, x_1)]$ . This composition of functions can be thought of as a generalization of time invariant systems, where the state transition of  $y$  depends only on the current state, and not separately of time,  $t$ . On the other hand, when there is a  $t$  present in front of the  $F$  term, then the system is said to be time variable. So it is possible



to have either time variable or time invariant systems with or without memory. Composite systems are basically the type of continuous flow system, or partial output, systems investigated by Schweickert and colleagues (Schweickert, 1989, 1998), and in some of our results below.

The most general of continuous flow dynamic systems and those that are required of the great bulk of applications, demand a dependence of the transition functions over an entire time interval, say  $(t_1, t)$ .<sup>1</sup> This means that the notation and math must take into consideration input  $u(t)$  to  $X$  or  $X$  to  $Y$  over the entire temporal interval  $(t_1, t)$ . This leads the mathematical operations into so-called *functionals*. Recall from above that composite functions map spaces or functions in a point-to-point fashion. For instance, in the above, a particular value of  $u$  at time  $t$ ,  $u(t)$  is mapped directly to  $x(t)$ .

In contrast, functionals map functions over entire intervals from, say the  $U$  space to the  $X$  space, as in  $u(t)$  into other functions (e.g.,  $x(t)$  over the interval  $(t_1, t)$ ). Thus, the value of  $x(t)$  will be a function of that entire interval, from  $t_1$  to  $t$ , not just that at  $t$ . In many cases, this dependence is due to a feedback loop from the output (e.g.,  $x(t)$ ) back into the input (e.g., perhaps added to an external input,  $u(t)$ ). In general, there is a mathematical equivalence between memory due to a feedback loop and a time-dependence that is due to, for instance, system aging. This equivalence can manifest itself as a quantitative equivalence between a closed-loop system and an open loop system. The global flow transition functions are easily expanded to get

$$F : T \times T \times X \times X \times U \rightarrow X$$

The time has come to introduce differential equations which are arguably one of the most valuable tools in all areas of engineering and science. Science has made considerable progress through differential equations. These express the rates of change of the state variables over time, as functions of the several variables and parameters of the system. In the classical theory of ordinary differential equations, weak regularity assumptions lead to a unique global description following from a specific local description (Grimshaw, 1990; Sanchez, 1979). Nevertheless, the latter may not be determined in closed form, that is expressible in easily understood, and compactly written functions. For instance, even in linear systems which are time variable (i.e., the system “ages” so that the same input with the same initial conditions but, at two different times can have extremely different effects), this can occur. More exotic examples are obviously encountered in chaotic systems (e.g., Devaney, 2018;

<sup>1</sup>In the technicalities of mathematical operations in systems theory, it can be critical to discern between, say, open intervals such as  $(t_1, t)$  and closed intervals like  $[t_1, t]$  but we can with impunity, avoid such niceties here.

Townsend, 1990). Hence, for this and other reasons, the differential equations form an axis of theoretical interest. Even when no closed form is available, the differential equations can usually be approximated by difference equations and results computed on digital computers.

Continuing with our notation, we let the instantaneous transition be designated by small  $f$  to show how the rate of change of state occurs. We pause to observe that our  $F$  and  $G$  are produced by integrating the derivatives  $\frac{d}{dt}x(t)$  and  $\frac{d}{dt}y(t)$  respectively. As before, the more global  $F$  and  $G$  are often referred to in the literature on dynamic systems as *flow* as opposed to the generating differential equations. We shall return to the former more global description presently.

Written in a local version as a differential equation we have,

$$\frac{d}{dt}x(t) = f(t, u, x)$$

The presence of  $t$  in the parenthesis means that, in general, the system can be time variable, that is, nonautonomous. As intimated above, for instance, the system might fatigue or even warm-up as a function of time. Further, note that because we are dealing with instantaneous transitions here, we do not require an entire time interval for the expression. As we will see, this facet does not imply that the integral, resulting in the explicit mapping  $F$ , as before, will not be a functional.

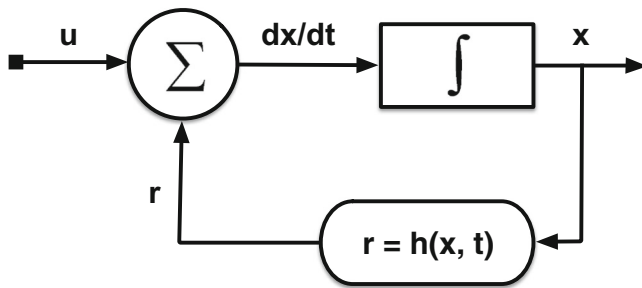
A solution to this type of differential system must involve integration to reverse the differentiation. Under some regularity conditions, one can find a solution by taking an approximation to whatever the real solution is and keep iterating the integration until the true solution is approached asymptotically. In fact, that is how it was originally proven that such a differential equation has one and only one solution (Sanchez, 1979). The presence of  $u$  means that the rate of change is a function of some input as well. And,  $x$  itself implies that the rate also depends on the state at exactly time  $t$  and this in itself suggests that the  $X$  subsystem can be expressed as possessing a feedback loop. Figure 3 shows the important case of a linear system with dependence on  $u(t)$ ,  $t$  and  $x$ . This particular type of system will be detailed below, but it does serve to reveal the types of critical aspects of time-invariant, linear systems.

Next, if

$$\frac{d}{dt}x(t) = f(u, x)$$

the system is said to be autonomous because the subsystem does not change inherently over time, so that the changes of state,  $x$ , are only dependent on the current state of the system and the input  $u$ .

If  $u = 0$  for all time  $t > t_1$ , then we are observing a zero-input system. Such a system might be able to deliver non-trivial output due to its initial state and memory or



**Fig. 3** Internal feedback through an integrator, illustrating a model of a system with memory. If the transformation  $h()$  on the “fed back” state  $x$  does not depend explicitly on time and is linear and sign inverting (i.e.,  $r = -ax$ ), this represents the familiar negative feedback loop

simply nonlinear characteristics. For instance, nonlinear systems can seem to create or pump in, new energy from no obvious apparent source.. In this case, we can put down the governing equation as

$$\frac{d}{dt}x(t) = f(t, x)$$

Inputs to a system are often defined by adding on a function of time to the terms involving only the state variable and its derivatives. If the system possesses no internal feedback or other innate time dependence, the  $x$  is excluded on the right-hand side, so that

$$\frac{d}{dt}x(t) = f(t, u)$$

From a more abstract standpoint, the only way in which a subsystem, say  $X$ , can exhibit its system characteristics through the memory in a system is by  $X$  being nondegenerate in the domain of the functional  $F$ . That is, past values of the input  $u$  can only affect the present transition by way of the memory of  $X$ . But this is again simply the feedback operation shown in Fig. 3.

In general, the functionals in a system take functions into numbers. That is,  $F$  and  $G$  are typically not true functions, because the latter take numbers into numbers. However, a special case ensues when the subsystems associated with subsystems  $X$  and  $Y$  are limited to an infinitesimal time point (i.e.,  $t_2 = t_1$ ). However, the possibility of a strategic role for initial conditions as the interplay between behavior and neuroscience evolves, should not be overlooked. Moreover, much the same can be said with regard to the input function  $u$ .

We now specialize our account to linear systems. Most readers are familiar with linearity:  $F$  is a linear function of its inputs if and only if  $F(u_1 + u_2) = F(u_1) + F(u_2)$  and  $F(au) = aF(u)$  for every constant  $a$ , and similarly for the second,  $Y$  stage. Operations like multiplication, taking the logarithm, taking the powers (e.g., quadratic as in  $x(t)^2$ , etc.) move a system into the realm of nonlinearity.

For instance, Grossberg’s (1988) theoretical approach has utilized nonlinear elements, and especially multiplication, to good advantage.

When we study a zero-input linear system it is called homogeneous, otherwise it is non-homogeneous (e.g., Luenberger, 1979; Padulo & Arbib, 1974) When a system is linear with an input, the transition state mapping and therefore the flow, is given by a weighted integral, where the weight function of the subsystem is called the *impulse response function* called  $h(t)$ . This system does possess memory, except in degenerate circumstances and therefore Figure 3 would retain the “T” symbol. The impulse response function  $h$  determines the importance that the input from different points in time in the past has on the current state, quite manifestly an indicant of memory (e.g., Townsend & Ashby, 1983, pp. 401-409). The name comes from the fact that if the input occurs at a single point (e.g., an impulse or spike; mathematically, a Dirac delta function; see below), then the system’s response,  $x(t)$ , equals  $h(t)$ . A special case of some interest is where the impulse response function is exponential in form, demonstrating a decreasing importance of input events as they recede into the past. However, many systems possess “memories” that are first increasing, then decreasing, functions of time past (Busey & Loftus, 1994).

Again, if a linear system is time variable and has an input, then  $h$  is a function of the present time  $t$ , as well as the past time  $t'$ , and the linear global state transition operator is given by

$$x(t) = \int_{t_1}^t h(t, t')u(t')dt'$$

This term is referred to as the *forced-function solution* because the output at time  $t$  in this term depends on (and only on) what the input is from  $t = 0$  to time  $t$ . However, if the linear differential system is time invariant, it is called homogeneous and by definition has no input, and the output is the single term  $x(t_1)$  which is, of course, just the initial condition and the solution is referred to as the *free response* or *free solution*.

When we combine both the free and forced parts we receive

$$x(t) = x(t_1) + \int_{t_1}^t h(t, t')u(t')dt'$$

Notice the summation of the term  $x(t_1)$  and the integral involving the input. This is a standard property of linear systems. Moreover, we can immediately see the linearity of such a system by observing the result of sending in the

sum of two inputs  $u(t) + v(t)$  (and ignoring any initial conditions):

$$x(t) = \int_{t_1}^t h(t, t')[u(t') + v(t')]dt'$$

However, if the basic system (ignoring input as a function of time) is autonomous, then the state transition depends only on the difference between the present time and the past time, not the actual value of the past time, so the impulse response function  $h(t, t') = h(t - t')$  replaces the general nonautonomous term in the above equation (e.g., see Townsend & Ashby, 1983, pp. 401-409). It can be seen that when the dummy variable  $t'$  is equal to  $t$ , the present,  $h(0)$  appropriately gives the weight when no memory is required, that is we are at the present instant of time. Assuming the input begins at  $t_1 = t$  then when  $t' = 0$ , the memory is, aptly,  $t$  time units long.

Retreating for a moment back to possibly non-linear composite systems we confront the fact that they are so elementary that one would rarely find such a process in natural phenomena described by way of differential equations. But, for completeness and assuming differentiability, the sister differential equations would be

$$\frac{d}{dt}x(t) = \frac{d}{dt}F(t) = f(t)$$

and

$$\frac{d}{dt}y(t) = \frac{d}{dx}G(t) \frac{d}{dt}x(t).$$

And the comparable flow is given by

$$x(t) = F(t) = \int_0^t f(t')dt'$$

and then

$$y(x(t)) = G(F(t)) = \int_0^t \left(\frac{d}{dx}G(t)\right) \left(\frac{d}{dt}x(t)\right) dt$$

In a way, this expression is both trivial and complex since we don't have any further information about the functions.

We can next incorporate an input function  $u(t)$  after  $t$  and before  $F$ . Written in a way that readers may find even more intuitive, we can now express the first stage as

$$\frac{d}{dt}x(t) = f(t, u)$$

and the second as

$$\frac{d}{dt}y(t) = g(x).$$

More complexity can be had by including an initial starting state but we shall move on at this time. If a system is time invariant but depends on current state (i.e., has memory) and possesses some input, then we find

$$\frac{d}{dt}x(t) = f(x, u).$$

If it is time-variable with memory, but with no input, we write

$$\frac{d}{dt}x(t) = f(t, x),$$

and so on. Analogous expressions can be written for the second stage.

A natural way of producing a memoryless, composite linear system is to let the impulse response function degenerate into something proportional to a so-called Dirac delta function,  $\delta(t)$ , which we can loosely define as a function, which is zero everywhere but at zero, where it integrates to 1.<sup>2</sup> That is,

$$\int_{-t}^{+t} \delta(t')dt'.$$

The systemic meaning is that the subsystem now has no memory: Only the present instantaneous input is passed through the system, weighted by a coefficient that may (in a non-autonomous system), or may not (autonomous system) depend on time. For instance, in an autonomous system, we have

$$x(t) = \int_{t_1}^t v(t')\delta(t - t')u(t')dt' = v(t)u(t),$$

a linearly-weighted function of the input  $u$ .

We show next how the concepts that we have developed can be used to solve simple linear differential equations. Our space is far too limited to laboriously work out all the details for the more complex models but we refer the reader to standard sources such as Padulo and Arbib (1974) and Luenberger (1979) and a number of other sources.

First, we justify the rubric for the impulse response function by inserting the Dirac delta function (the mathematical version of an impulse or spike) as input into a forced response solution, which weights the input by the memory (the  $\delta(t)$  function):

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

which comes about because  $\delta(t') = 0$  when and only when  $t' = t$ .

<sup>2</sup>A rigorous development of  $\delta(t)$  requires the notion of generalized functions or, equivalently, so-called distributions (Gelfand & Shilov, 1964).



Even though we have to eschew most of the technicalities associated with present developments, we pause to motivate our larger points with a very simple case. Consider a homogeneous, time invariant, linear system, with an input but no initial condition, starting at  $t' = 0$ . Our differential equation (which can be taken as describing our first subsystem) is then

$$\frac{d}{dt}x(t) = -ax(t) + u(t)$$

Our solution for the impulse response function is

$$h(t) = e^{(-at)}$$

recalling that this time the  $t$  stands for the time difference from  $t' = 0$  to  $t' = t$ . We prove this claim by inserting the proposed  $h$  function in the usual integral:

$$\begin{aligned} x(t) &= \int_0^t h(t-t')u(t')dt' \\ &= \int_0^t e^{[-a(t-t')] }u(t')dt' \end{aligned}$$

Taking the derivative we get

$$\begin{aligned} \frac{d}{dt}x(t) &= \int_0^t (-a)e^{[-a(t-t')] }u(t')dt + u(t) \\ &= -ax(t) + u(t) \end{aligned}$$

This style of development can be employed to derive the more general results we desire but fortunately, Pierre Simon-Laplace, born a couple of decades after the passing of Isaac Newton, gave us the useful Laplace transform. It relies on the fact that most functions we are likely to meet possess an infinite number of guises, one obtained from the other via mathematical transformation. A desirable property of such a transform, and one claimed by the Laplace transform, is that the mapping from one representation to the other is 1-1, that is the Laplace transform of a well-behaved function of time is unique, and vice versa. In this case,

$$\mathcal{L}\{f(t)\} = \mathbf{f}(s)$$

Notice that the function will itself be different and a function, now, of the so-called transform variable  $s$ . This transform possesses many useful aspects. One is that the

transform of a derivative of  $f(t)$  is just  $\mathbf{f}(s)$  itself multiplied by  $s$ .

Now recall that the way in which memory expresses itself in linear systems is by way of the convolution of  $h$  and the input  $u(t)$ . That is,

$$\begin{aligned} x(t) &= \int_0^t h(t-t')u(t')dt' \\ &= \int_0^t e^{[-a(t-t')] }u(t')dt' \end{aligned}$$

A second property of the transform is that it converts the unwieldy convolution to a product in the  $s$ -space. Thus,

$$u(t) * h(t) : \mathcal{L} \rightarrow \mathbf{u}(s) \times \mathbf{h}(s).$$

Using these aspects of the transform to good advantage we find that

$$\frac{d}{dt}x(t) = -ax + u(t) : \mathcal{L} \rightarrow \mathbf{s}\mathbf{x}(s) = -\mathbf{a}\mathbf{x}(s) + \mathbf{u}(s).$$

So, with simple algebra, we solve for  $\mathbf{x}(s)$  to achieve

$$\mathbf{x}(s) = \frac{\mathbf{u}(s)}{\mathbf{a} + s}$$

Now we can utilize the fact that the inverse Laplace transform,  $\mathcal{L}^{-1}$  of the right hand side is the convolution where we find that  $h(t)$  is the exponential that we just earlier proved through the other method. This is because the inverse transform of  $1/(a + s)$  is just  $e^{-(at)}$ , and we are done. We have thus shown two independent ways of finding out what the convolution in a very simple linear system is.

## Applying the analysis: generalizing the cascade model

We now use these tools to solve for a generalized version of the cascade model (McClelland, 1979). In order to render a fairly complex model into something that can be readily comprehended, we take up the special case of two parallel channels. We will call the first the signal or target channel (channel I) and will call the other the flanker channel (channel II). Each of these pathways proceeds in not only a continuous flow fashion, but the overall system is governed by homogeneous (i.e., time-invariant) linear differential equations with two sequential subprocesses.

The original cascade model ends with a decision criterion that is fixed within a trial. When met and exceeded, the system entails a decision and consequent response. Because the rest of the system is deterministic, in order to introduce some modest stochasticity into the action,

that criterion is assumed to vary from trial to trial with a normal distribution. One of the first to employ this device was Clark Hull 1952, the famous animal neo-behaviorist. More within the realm of human cognitive psychology, Grice and colleagues (Grice, Nullmeyer, & Spiker, 1977) also modeled data with this construction. Dzhafarov (1993) demonstrated the mathematical equivalence, once again evoking the specter in the behavioral sciences of model mimicking of a set of models obeying this principle vs. a set with fixed criterion but stochastic activations (see, e.g., Townsend and Ashby, 1983, chapter 14).

In order to fulfill Eriksen and company’s hypothesis that both input streams, the signal’s and the flanker’s, appear to influence even the arm’s trajectory through the lift-off of the hand to the final button press, we must reinterpret the decision criterion as the final touch-down of the finger. Figure 4 which represents as a flow diagram our cascaded, continuous flow model for the flanker task. As in our earlier general treatment  $x_i$  will stand for the first sub-process accepting input with  $i = 1, 2$ , and  $y_j$  will represent the next sub-process. We depict the cascade model as using feedback loops circling back to sum with the input. This depiction is therefore called a “closed loop system” in the literature.

Readers may wish to try their hand at drawing a flow-diagram using the language employed by McClelland 1979 before turning to Fig. 4. His portrayal emphasizes that each channel can be interpreted as “open loop” control subsystems where the input is driving the state (i.e.,  $x$  or  $y$ ) toward the input level, in a way that is proportional to the difference between the input and the current level (of  $x$  or  $y$ ). In contrast, our closed loop portrayal reveals how feedback control systems work. Although the two guises are mathematically equivalent, in actual physical

construction of dynamic systems, the feedback control architecture provides more stable systems.

The differential equations pertaining to the first level (i.e., the  $x$  level) are

$$\frac{d}{dt}x_1(t) = k_{x_1} (u_1(t) + w_{2,1}(0)u_2(t) - x_1(t)) \tag{1}$$

$$\frac{d}{dt}x_2(t) = k_{x_2} (u_2(t) + w_{1,2}(0)u_1(t) - x_2(t)) \tag{2}$$

and those governing the second ( $y$ ) level are

$$\frac{d}{dt}y_1(t) = k_{y_1} (x_1(t) + w_{2,1}(1)x_2(t) - y_1(t)) \tag{3}$$

$$\frac{d}{dt}y_2(t) = k_{y_2} (x_2(t) + w_{1,2}(1)x_1(t) - y_2(t)) \tag{4}$$

The notations in Fig. 4 are intended to aid interpretation the various parameters in Eqs. 1-4. We let  $w_{i,j}(k)$  stand for the weight of channel  $i$  (associated with level  $k$  on the receptor (node, etc.) in the next higher level  $j$ . Thus,  $w_{2,1}(0)$  is the weight at the 0th level (i.e., that of the input) on the input from channel II (crossing) over to Channel I at level 1.

Note that we are arbitrarily setting what would be, say,  $w_{1,1}(0)$  to the value 1 because this can simply be considered to scale the various parameters and variables. That is why  $w_{1,1}(0)$  doesn’t actually appear in the figure. Observe that this structure permits forward cross-talk from one channel to another, perhaps facilitatory, perhaps inhibitory. Also, observe that the current value of the state is subtracted from the overall input. The reader will now infer that, for instance,  $w_{1,2}(1)$  tells us the weight from the  $x$  stage and

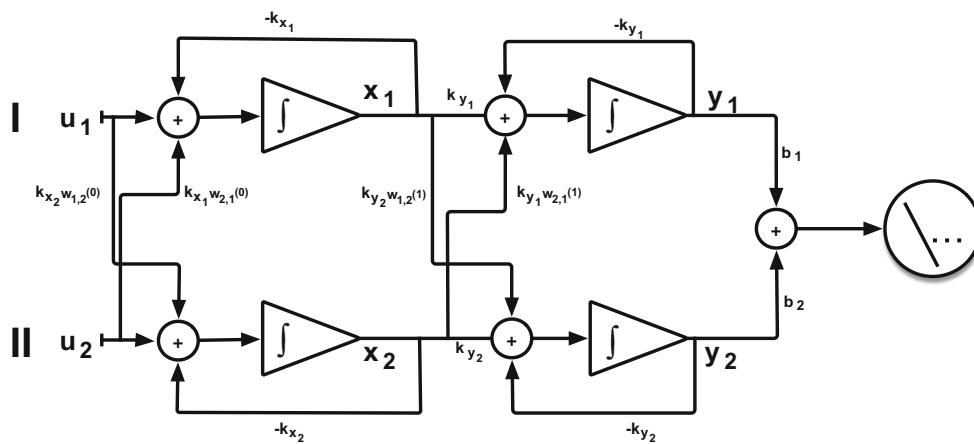


Fig. 4 A continuous flow instantiation of the cascade model. The small notations to the side of each element indicate multiplicative weights

channel I (by definition, now at level 1) to channel II at the  $y$  stage (level 2).

Next,  $k_{i,j}$  where  $i = x, y$  and  $j = 1, 2$  serve as rates of change of channels 1 or 2 at either the  $x$  stage or the  $y$  stage. At this one level, our version is more general than McClelland's. He assumed that  $k_{i,1} = k_{i,2}$ , for  $i = x$  or  $y$ . This implies that at level  $x$ , the gain or rate of change to channel I was the same as for channel II and similarly for the  $y$  level or sub-process. This means that although the  $w'$ s allowed  $u_1$  to affect  $x_1$  differently than it did  $x_2$  the rate of change at the next level had to be the same for both channels. Our rendition is more general in this sense but also naturally, more cumbersome.

In order to allow easier comparison with McClelland's (1979) matrix formulations, Eqs. 5–6 place our key differential expressions in a vector and matrix format. The arrows then directly picture the Laplace transform of these

time-descriptions over to the transform  $s$  space, also in vector-matrix notation.

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = (k_{y_1} \ k_{y_2}) \begin{bmatrix} x_1(t) + w_{2,1}(1)x_2(t) - y_1(t) \\ x_2(t) + w_{1,2}(1)x_1(t) - y_2(t) \end{bmatrix} \quad (5)$$

$\rightarrow \mathcal{L}$

$$s \begin{pmatrix} y_1(s) \\ y_2(s) \end{pmatrix} = (k_{y_1} \ k_{y_2}) \begin{bmatrix} x_1(s) + w_{2,1}(1)x_2(s) - y_1(s) \\ x_2(s) + w_{1,2}(1)x_1(s) - y_2(s) \end{bmatrix} \quad (6)$$

The next expressions picture the  $s$  functions, having solved for each  $\mathbf{x}(s)$  and each  $\mathbf{y}(s)$  in terms of the inputs, weights, and  $s$  itself. Then, in order to acquire a representation of the second  $y$  level in terms of everything before, we must substitute the solution for the  $x$ s into that of the  $y$ s, giving us

$$y_1(s) = \left\{ k_{x_1} \frac{u_1(s) + w_{2,1}(0)u_2(s)}{(s + k_{x_1})(s + k_{y_1})} + w_{2,1}(1)k_{x_2} \frac{u_2(s) + w_{1,2}(0)u_1(s)}{(s + k_{x_2})(s + k_{y_1})} \right\} k_{y_1} \quad (7)$$

$$y_2(s) = \left\{ k_{x_2} \frac{u_2(s) + w_{1,2}(0)u_1(s)}{(s + k_{x_2})(s + k_{y_2})} + w_{1,2}(1)k_{x_1} \frac{u_1(s) + w_{2,1}(0)u_2(s)}{(s + k_{x_1})(s + k_{y_2})} \right\} k_{y_2} \quad (8)$$

At this point, we are getting close to where greater ease will transpire in our operations if we make the simplifying assumption (following McClelland, 1979) that the input  $u$  functions are step functions. He assumed that  $u(t)$  was 0 up until  $t = 0$  and then  $u(t) = 1$  in either channel and this input was assumed to remain on until a decisional criterion was reached. This implied that the  $u_1$  information (e.g.,

feature, etc.) was coded the same as the  $u_2$  information. This restriction is readily obviated by letting  $u_1(t)$  not be equal to  $u_2(t)$  but both are, like McClelland's input, step functions:  $u_1(t) = u_1$ , a constant and the same for  $u_2(t) = u_2$ .

We can now (following McClelland, 1979) bring to bear what are known as partial fraction operations that will make it straightforward to apply inverse Laplace transforms.<sup>3</sup> The partial fraction decompositions are expressed as

$$y_1(s) = \left( \frac{u_1}{s} + w_{2,1}(0) \frac{u_2}{s} \right) \frac{k_{x_1}k_{y_1}}{(s + k_{x_1})(s + k_{y_1})} + w_{2,1}(1) \left( \frac{u_2}{s} + w_{1,2}(0) \frac{u_1}{s} \right) \frac{k_{y_1}k_{x_2}}{(s + k_{y_1})(s + k_{x_2})} \quad (9)$$

$$y_2(s) = \left( \frac{u_2}{s} + w_{1,2}(0) \frac{u_1}{s} \right) \frac{k_{y_2}k_{x_2}}{(s + k_{y_2})(s + k_{x_2})} + w_{1,2}(1) \left( \frac{u_1}{s} + w_{2,1}(0) \frac{u_2}{s} \right) \frac{k_{y_2}k_{x_1}}{(s + k_{y_2})(s + k_{x_1})} \quad (10)$$

$$y_1(s) = (u_1 + w_{2,1}(0)u_2) \left[ \frac{1}{s} - \frac{1}{s + k_{x_1}} \frac{k_{y_1}}{k_{y_1} - k_{x_1}} - \frac{1}{s + k_{y_1}} \frac{k_{x_1}}{k_{x_1} - k_{y_1}} \right] + w_{2,1}(1)(u_2 + w_{1,2}(0)u_1) \left[ \frac{1}{s} - \frac{1}{s + k_{x_2}} \frac{k_{y_1}}{k_{y_1} - k_{x_2}} - \frac{1}{s + k_{y_1}} \frac{k_{x_2}}{k_{x_2} - k_{y_1}} \right] \quad (11)$$

$$y_2(s) = (u_2 + w_{1,2}(0)u_1) \left[ \frac{1}{s} - \frac{1}{s + k_{x_2}} \frac{k_{y_2}}{k_{y_2} - k_{x_2}} - \frac{1}{s + k_{y_2}} \frac{k_{x_2}}{k_{x_2} - k_{y_2}} \right] + w_{1,2}(1)(u_1 + w_{2,1}(0)u_2) \left[ \frac{1}{s} - \frac{1}{s + k_{x_1}} \frac{k_{y_2}}{k_{y_2} - k_{x_1}} - \frac{1}{s + k_{y_2}} \frac{k_{x_1}}{k_{x_1} - k_{y_2}} \right] \quad (12)$$

<sup>3</sup>The are many expositions on partial fractions, but a quick and clear description can be found in Wikipedia.

And, the succeeding inverse transforms are readily found to be

$$\begin{aligned} \mathcal{L}^{-1} \rightarrow \\ y_1(t) &= (u_1 + w_{2,1}(0)u_2) \left[ 1 - \frac{k_{y_1}}{k_{y_1} - k_{x_1}} e^{-k_{x_1}t} - \frac{k_{x_1}}{k_{x_1} - k_{y_1}} e^{-k_{y_1}t} \right] \\ &\quad + w_{2,1}(1)(u_2 + w_{1,2}(0)u_1) \left[ 1 - \frac{k_{y_1}}{k_{y_1} - k_{x_2}} e^{-k_{x_2}t} - \frac{k_{x_2}}{k_{x_2} - k_{y_1}} e^{-k_{y_1}t} \right] \end{aligned} \tag{13}$$

$$\begin{aligned} y_2(t) &= (u_2 + w_{1,2}(0)u_1) \left[ 1 - \frac{k_{y_2}}{k_{y_2} - k_{x_2}} e^{-k_{x_2}t} - \frac{k_{x_2}}{k_{x_2} - k_{y_2}} e^{-k_{y_2}t} \right] \\ &\quad + w_{1,2}(1)(u_1 + w_{2,1}(0)u_2) \left[ 1 - \frac{k_{y_2}}{k_{y_2} - k_{x_1}} e^{-k_{x_1}t} - \frac{k_{x_1}}{k_{x_1} - k_{y_2}} e^{-k_{y_2}t} \right] \end{aligned} \tag{14}$$

These, then, turn the solutions of the differential cascade system equations into the real-time functions. We can see that, due to the assumptions of time-invariance and also of the elementary nature of the inputs, the final results have essentially already integrated the convolutions and delivered the consequent weighted exponential functions.

Ordinarily, at this point we would append decisional (e.g., detection, classification, etc.) criteria in the channels I, II with which  $y_1(t)$  and  $y_2(t)$  are compared on a continuing basis. Then a final decision and response would be implemented. For instance, when  $y_1(t)$  and/or  $y_2(t)$  reaches its criterion it will provoke a response (e.g., a 0 or 1 output in each channel) which might then input to a logic gate, for instance an OR or AND gate, in order to arrive at the final decision and response. We have constructed and implemented such models in many studies over the past sixteen years (e.g., Eidels et al., 2011; Townsend & Wenger, 2004, 2012; Wenger & Townsend, 2006). However, in the present context, the observer is supposed to be ignoring the activation in the flanker channel (channel II). Thus, we treat the signal channel (Channel I) as leading to the continuous, indicator-like response (representing the wavering, in-transit finger motion) with the flanker information affecting that trajectory in either a facilitatory or inhibitory manner.

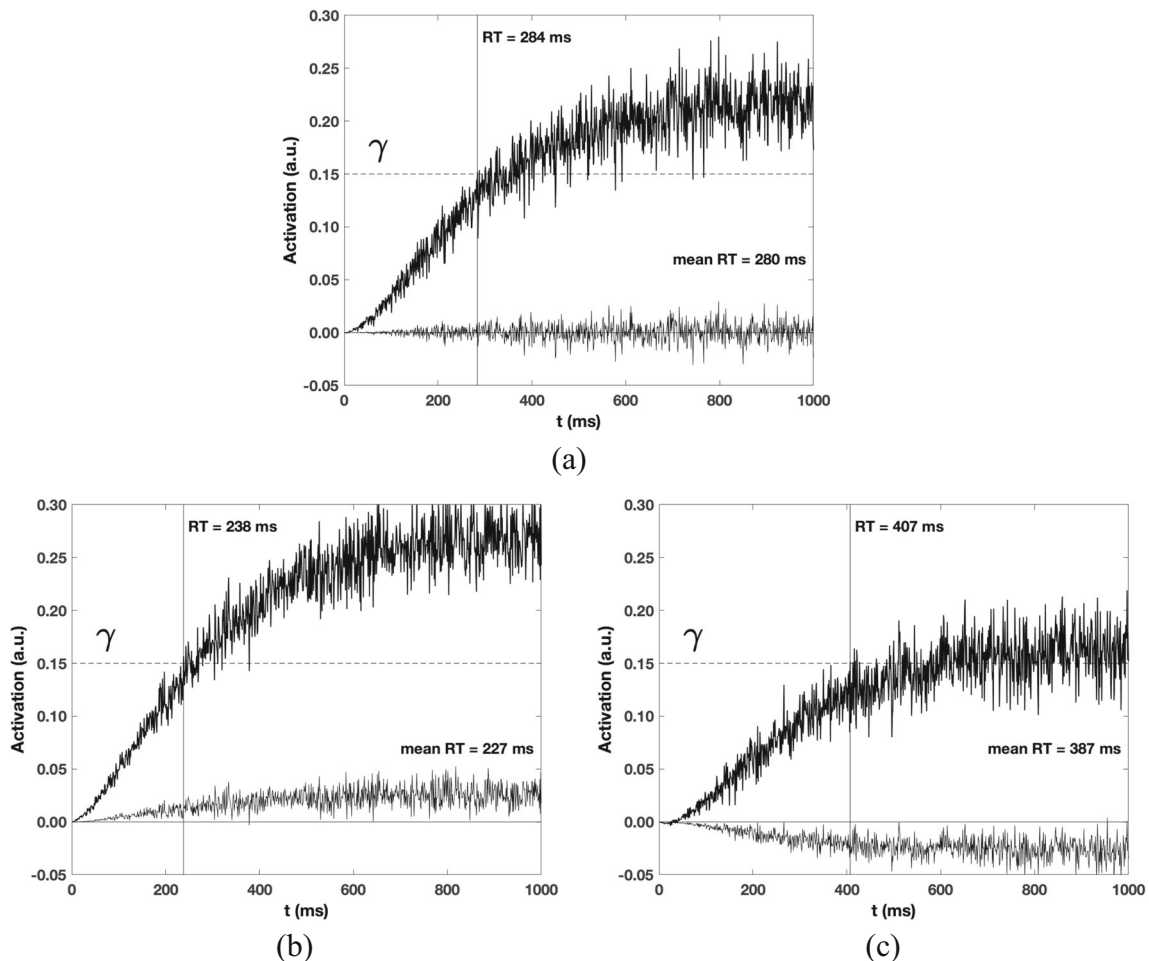
In many flanker designs, there are two signals or signal classes. In the present, exemplary modeling exercise, we assume there is no substantive difference. We thereby institute a vector dot product of  $(y_1(t), y_2(t))$  by  $(b_1, b_2)$  into  $z(t) = (y_1(t), y_2(t)) \cdot (b_1, b_2)$  which effectively provides a weighted sum of the  $y$  outputs into the wavering finger (i.e.,  $z$ ). The  $b_1$  entry can be assumed to always weight the information coming in the signal channel (i.e., Channel I) and let  $b_2$  be, for example, a positive number, for example,  $b_2 = +b$  if the flanker (in Channel II) agrees with the presented signal information but  $b_2 = -b$  if the flanker is similar to the other signal. We pause to observe

that this operation is similar to that of coactivation systems (e.g., Townsend & Wenger 2004) where two channels feed their combined activations into a final conduit. Usually, this operation is a sum of two positive activations and is typically unweighted. Therefore, the present model can be viewed as a mild generalization of the coactivation concept. Lastly, we can capture the ‘touchdown’ of the finger on a signal 1 or signal 2 location by assuming that occurs when  $z(t) = +r_1$  or conversely  $z(t) = -r_2$ .

As in our other efforts modeling parallel systems, noise should be added in applications. The ensuing mathematics lie within rather abstruse regions of very general stochastic differential equations, that can be extremely challenging to solve analytically (e.g., for the first passage times), and such concerns go far beyond our present aims. For illustration, we have carried out a small set of simulations with this model. Figure 5a shows a simulation results in the situation where the flanker is neutral, Fig. 5b shows the result when the flanker is consistent, and Fig. 5c shows the result when the flanker is inconsistent. As can be seen in these three panels, the predicted RTs are ordered as expected.

### The generalities and constraints of this account

The systems analyzed here were limited to a set of two sequentially linked, subsystems. Table 1 includes Dondersian systems because we think it is important to reveal how a general continuous flow dynamic system can be rigorously morphed into a Dondersian system. However, after that, the latter are of no further concern to us here. All our subsequent systems are lumped, meaning that, informally, input to the first subsystem has an instantaneous effect that is instantaneously transmitted to the second subsystem. This element is most critical for representing the kind of dynamic presaged by the Eriksen and colleagues’ efforts. Our covered systems include no systems based on partial differential equations nor any consideration of



**Fig. 5** Example activations for one trial in each of the three typical conditions of the flanker task: **a** neutral flankers, **b** consistent flankers, **c** inconsistent flankers. The *upper trace* in each panel shows the combined activation of the target and flanker channels, and the *bottom*

*trace* shows the activation for the flanker channel alone. The choice threshold is labeled as  $\gamma$ , and the RT on the trial marked by the *vertical reference line*. The mean RTs for that condition (from 100 simulated trials) is also noted in each panel

time-lags. Most of the complex systems encompassed in the OPNETs of Goldstein and Fisher (1991) or even PERT networks of Schweickert, Fisher, and Sung (2012) are excluded. Although feedback is allowed within-subsystem as a mark of dynamic stability, we eschew it across subsystems. This is a limitation that will certainly be set aside in future efforts given its obvious value in neural and behavioral dynamics.

The most severe constraint on our taxonomy, in terms of modern cognitive modeling, is the absence of stochasticity. Most of our theorems and predictions should be valid in the mean but ultimately researchers must turn to stochastic systems. The best studied continuous state and time processes are founded on strong Markov assumptions and Martingale properties (Karlin & Taylor, 1981; Smith & Van Zandt, 2000) but still of very high mathematical complexity, will be the class of diffusion processes. In particular, the drift-diffusion (Wiener) process developed and explored by many years by Ratcliff and colleagues (beginning with

Ratcliff, 1978), took some time to take hold in the field. That is, though one of the simplest diffusion systems, it is still much more complicated than simpler models, like the linear ballistic accumulator model of Brown and Heathcote (2008) or even models with, say, discrete time but continuous state space as in the models explored by Link and Heath (1975) or discrete state but continuous time via the accumulator model of Smith and Vickers (1989). The exclusion of a huge set of dynamic systems, permits a deep and general penetration of sequential, lumped systems and, we hope, provides the potential for furthering the kind of theorizing foreseen by Eriksen.

With respect to connecting theory and data, two natural next steps seem apparent. The first is to broaden the set of measures that could be used to assess response competition in continuous flow models. This could include those discussed in the present collection by Erb et al. 2020, the use of force transducers (Ray, Slobounov, Mordkoff, Johnston, & Simon, 2000), or potential modifications to the lateralized



readiness potential (e.g., Mordkoff & Grosjean, 2001; Wenger & Rhoten, 2020). The second step, pertinent to any of the possible dependent variables, concerns predictions for factorial interactions in the observed dependent variables (Schweickert, 1978; Sternberg, 1969; Townsend & Ashby, 1983). A set of theory-driven methodologies has arisen from such questions referred to as *systems factorial technology* (Little, Altieri, Yang, & Fific, 2017; Townsend & Nozawa, 1995). One critical key of systems functioning is whether the influence of two or more factors is additive in the dependent variable. The classical dependent variable has been response time (Sternberg, 1969; Townsend & Ashby, 1983). Foundational work by Schweickert (1978, 1985) established new results for response times but also extended such analyses to states of the system. This type of exploration is quite challenging because he considered completely general composite systems of the type found in our above taxonomy. For that reason, the targeted systems were deterministic as in our taxonomy, but should be relevant to expected (mean) values of the dependent variables. One of our goals is to extend such deep analyses to memorized continuous flow systems, again first confining ourselves to deterministic systems.

## References

- Anderson, J. A. (1995). *An introduction to neural networks*. Cambridge: MIT.
- Ashby, R. (1952). *Design for a brain: The origin of adaptive behavior*. London: Chapman and Hall.
- Ashby, F. G., & Townsend, J. T. (1980). Decomposing the reaction time distribution: Pure insertion and selective influence revisited. *Journal of Mathematical Psychology*, *21*, 93–123.
- Ashby, F. G. (1982). Deriving exact predictions from the cascade model. *Psychological Review*, *89*, 599–607.
- Bingham, G. P., Schmidt, R. C., & Rosenblum, L. D. (1995). Dynamics and the orientation of kinematic forms in visual event recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1473.
- Booth, T. L. (1967). *Sequential machines and automata theory*. Wiley.
- Boring, E. G. (1957). *History of experimental psychology*. New York: Appleton-Century-Crofts.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178.
- Busemeyer, J. B., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.
- Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review*, *101*, 446–469.
- Coles, M., Gratton, G., Bashore, T., Eriksen, C., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 529.
- Cunningham, W. J. (1958). *Introduction to nonlinear analysis*. New York: McGraw-Hill Book Co.
- Devaney, R. (2018). *An introduction to chaotic dynamical systems*. CRC Press.
- Donders, F. C. (1868/1969). Over die snelheid van psychische processen (trans. W. G. Koster). *Acta Psychologica*, *30*, 412–431.
- Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika*, *58*, 281–314.
- Eidels, A., Hout, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of Mathematical Psychology*, *55*(2), 176–190.
- Erb, C., Smith, K., & Moher, J. (2020). Tracking continuities in the flanker task: From continuous flow to movement trajectories. *Attention, Perception, & Psychophysics*. (in press).
- Eriksen, C. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics*, *12*(2), 201–204.
- Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics*, *14*(1), 155–160.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149.
- Eriksen, C., & Schulz, D. (1979). Information processing in visual search: Some theoretical considerations and experimental results. *Perception & Psychophysics*, *25*, 249–263.
- Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing diagrams. *Journal of Mathematical Psychology*, *27*, 121–151.
- Gelfand, I. M., & Shilov, G. E. (1964). *Generalized functions*, Vol. 1–4. Academic Press.
- Goldstein, W. M., & Fisher, D. L. (1991). Stochastic networks as models of cognition: Derivation of response time distributions using the order-of-processing method. *Journal of Mathematical Psychology*, *35*, 214–241.
- Grice, G. R., Nullmeyer, R., & Spiker, V. A. (1977). Application of variable criterion theory to choice reaction time. *Perception & Psychophysics*, *22*(5), 431–449.
- Grimshaw, R. (1990). *Nonlinear ordinary differential equations*. Oxford: Blackwell Scientific Publications.
- Grossberg, S. (1988). *Adaptive brain ii*. Elsevier.
- Haykin, S., & Fuster, J. M. (2014). On cognitive dynamic systems: Cognitive neuroscience and engineering learning from each other. *Proceedings of the IEEE*, *102*(4), 608–628.
- Heath, R. A., & Fulham, R. (1988). An adaptive filter model for recognition memory. *British Journal of Mathematical and Statistical Psychology*, *41*(1), 119–144.
- Hertz, J., Krogh, A., Palmer, R. G., & Hornet, H. (1991). Introduction to the theory of neural computation. *Physics Today*, *44*, 70.
- Hull, C. L. (1952). *A behavior system; an introduction to behavior theory concerning the individual organism*. Yale University Press.
- Karlin, S., & Taylor, H. E. (1981). *A second course in stochastic processes*. Elsevier.
- Klir, G. J. (1969). *Approach to general systems theory*. Van Nostrand Reinhold.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*, 77–105.
- Little, D., Altieri, N., Yang, C., & Fific, M. (2017). Systems factorial technology: A theory-driven methodology for the identification of perceptual and cognitive mechanisms. Elsevier Science & Technology Books. Retrieved from <https://books.google.com/books?id=OQICMQAACAAJ>.
- Luenberger, D. G. (1979). *Introduction to dynamic systems: Theory, models and applications*. New York: Wiley.

- MacDonald, N. (2013). *Time lags in biological models* (Vol. 27). Springer Science & Business Media.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 86, 287–330.
- Miller, J. (1988). Discrete and continuous models of human information processing: Theoretical distinctions and empirical results. *Acta Psychologica*, 67(3), 191–257.
- Molenaar, P. C., De Gooijer, J. G., & Schmitz, B. (1992). Dynamic factor analysis of nonstationary multivariate time series. *Psychometrika*, 57(3), 333–349.
- Mordkoff, J. T., & Grosjean, M. (2001). The lateralized readiness potential and response kinetics in response-time tasks. *Psychophysiology*, 38(5), 777–786.
- Padulo, L., & Arbib, M. (1974). *System theory: A unified approach to continuous and discrete systems*. Philadelphia: WB Saunders Company.
- Patterson, R. E., Fournier, L. R., Williams, L., Amann, R., Tripp, L. M., & Pierce, B. J. (2013). System dynamics modeling of sensory-driven decision priming. *Journal of Cognitive Engineering and Decision Making*, 7(1), 3–25.
- Port, R. F., & van Gelder, T. (1995). *Mind as motion: Explorations in the dynamics of cognition*. Cambridge: MIT Press.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ray, W., Slobounov, S., Mordkoff, J., Johnston, J., & Simon, R. (2000). Rate of force development and the lateralized readiness potential. *Psychophysiology*, 37(6), 757–765.
- Roberts, S., & Sternberg, S. (1993). The meaning of additive reaction time effects: Tests of three alternatives. In Meyer, D. E., & Kornblum, S. (Eds.) *Attention and performance xiv*, (pp. 611–654). Cambridge: MIT Press.
- Sanchez, D. A. (1979). *Ordinary differential equations and stability theory: An introduction*. Courier Corporation.
- Sanders, A. (1977). Structural and functional aspects of the reaction process. *Attention and performance*, 6, 3–25.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, 18, 105–139.
- Schweickert, R. (1985). Separable effects of factors on speed and accuracy: Memory scanning, lexical decision, and choice tasks. *Psychological Bulletin*, 97(3), 530.
- Schweickert, R. (1989). Separable effects of factors on activation functions in discrete and continuous models: D? and evoked potentials. *Psychological Bulletin*, 106, 318–328.
- Schweickert, R., & Mounts, J. (1998). Additive effects of factors on reaction time and evoked potentials in continuous-flow models. In Dowling, C. E., & Roberts, F. S. (Eds.) *Recent progress in mathematical psychology: Psychophysics, knowledge, representation, cognition, and measurement*, (pp. 311–327). Mahwah: Erlbaum.
- Schweickert, R., Fisher, D. L., & Sung, K. (2012). *Discovering cognitive architecture by selectively influencing mental processes* Vol. 4. World Scientific.
- Smith, P., & Vickers, D. (1989). Modeling evidence accumulation with partial loss in expanded judgment. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4), 797.
- Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, 53(2), 293–315.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of donders' method. *Acta Psychologica*, 30(0), 276–315.
- Taylor, D. A. (1976). Stage analysis of reaction time. *Psychological Bulletin*, 83, 161–191.
- Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of perception and action*. Cambridge: MIT Press.
- Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 25, 168–199.
- Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In Kantowitz, B. H. (Ed.) *Human information processing: Tutorials in performance and cognition*, (pp. 133–168). Hillsdale: Erlbaum.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: University Press.
- Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, 33, 309–327.
- Townsend, J. T. (1990). Truth and consequences of ordinal differences in statistical distributions: Toward a theory of hierarchical inference. *Psychological Bulletin*, 108, 551–567.
- Townsend, J. T., & Nozawa, G. (1995). On the spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, 39, 321–359.
- Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003–1035.
- Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology*, 56(6), 476–494.
- Van Zandt, T., & Townsend, J. T. (2012). *Mathematical psychology*. American Psychological Association.
- Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 755–779.
- Wenger, M. J., & Rhoten, S. E. (2020). Perceptual learning produces perceptual objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(3), 455–475.
- Zhang, Q., Walsh, M., & Anderson, J. (2018). The impact of inserting an additional mental process. *Computational Brain Behavior*, 1, 22–35.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.