BEHAVIORAL AND BRAIN SCIENCES (1998) **21,** 615–665
*Printed in the United States of America*

# The dynamical hypothesis in cognitive science

**Tim van Gelder**
*Department of Philosophy, University of Melbourne, Parkville VIC 3052, Australia*
**tgelder@ariel.unimelb.edu.au**
**ariel.its.unimelb.edu.au/~tgelder**

**Abstract**: According to the dominant computational approach in cognitive science, cognitive agents are digital computers; according to the alternative approach, they are dynamical systems. This target article attempts to articulate and support the dynamical hypothesis. The dynamical hypothesis has two major components: the nature hypothesis (cognitive agents *are* dynamical systems) and the knowledge hypothesis (cognitive agents can be understood dynamically). A wide range of objections to this hypothesis can be rebutted. The conclusion is that cognitive systems may well be dynamical systems, and only sustained empirical research in cognitive science will determine the extent to which that is true.

**Keywords:** cognition; computability; computational systems; computers; dynamical systems; modeling; systems; time

## 1. Introduction

Some five decades after *Principia mathematica,* David Hume dreamt of a scientific psychology in which mathematical laws would govern the mental realm, just as Newton's laws governed the material realm (Hume 1739–1740/1978). The universal force of gravitation, whereby bodies attract in proportion to their mass, would be replaced by a universal force of association, whereby ideas attract in proportion to their similarity. The dynamics of matter would be paralleled by a dynamics of mind.

The Humean dream was not the first vision of mind inspired by the emergence of modern science. The new physics had uncovered mathematical laws of great simplicity and elegance, but laborious calculation was required to derive the messy details of actual behaviors. Thomas Hobbes took this calculating activity itself as his model of the mechanisms of mental operation. Perhaps thought is symbolic computation, the rule-governed manipulation of symbols inside the head (Hobbes 1651/1962).

Seventeenth-century speculation became twentieth-century science. Hobbes's idea evolved into the *computational hypothesis* (CH) that cognitive agents are basically digital computers. Perhaps the most famous rendition is Newell and Simon's doctrine that "a physical symbol system has the necessary and sufficient means for general intelligent action." They proposed this hypothesis as a "law of qualitative structure," comparable to the cell doctrine in biology or plate tectonics in geology. It expresses the central insight of the research paradigm that has dominated cognitive science for some 40 years.

In recent years, however, the Humean alternative has been gaining momentum. One of the most notable developments has been the rise of connectionism, which models cognition as the behavior of dynamical systems (Smolensky 1988), and often understands those models from a dynamical perspective. Equally significant is the emergence of cognitive neuroscience, and within it, the increasing prevalence of dynamical theorizing. Dynamics forms the general framework for growing amounts of work in psychophysics, perception, motor control, developmental psychology, cognitive psychology, situated robotics and autonomous agents research, artificial intelligence, and social psychology. It is central to a number of general approaches, such as ecological psychology, synergetics, and morphodynamics.[1]

The *dynamical hypothesis* (DH) is the unifying essence of dynamical approaches to cognition. It is encapsulated in the simple slogan, *cognitive agents are dynamical systems.* The aims of this target article are (1) to *articulate* the hypothesis – that is, to explain what the slogan means – and (2) to *defend* it as an open empirical hypothesis standing as a substantive alternative to the CH. The DH contends for the status of the "law of qualitative structure" concerning the nature of cognition.

One goal in undertaking this philosophical work is to clarify the conceptual terrain. Another is to help clear rhetorical space for dynamicists in cognitive science to get on with



Tim van Gelder completed his Ph.D. at the University of Pittsburgh and then taught at Indiana University before taking up his current position as Queen Elizabeth II Research Fellow in the department of Philosophy at the University of Melbourne, Australia. He has published widely in the fields of philosophy of mind and the foundations of cognitive science, with particular focus on exploring alternatives to orthodox functionalist and computationalist perspectives.

© 1998 Cambridge University Press 0140-525X/98 $12.50

https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

the hard work of developing detailed accounts of specific cognitive phenomena. The most important goal, however, is to gain insight into the nature of people – for people are, among other things, cognitive agents.

This paper ploughs an interdisciplinary field. Boulders of ambiguity, vagueness, and confusion must be cleared away. Much effort is devoted simply to establishing a single coherent and reasonably precise framework for discussion. This framework involves commitments at terminological, conceptual, and even metaphysical levels. Its development requires many choices and stipulations, often somewhat arbitrary in nature. Occasional conflicts with existing intuitions are unavoidable. Still, *some* such regimentation is essential, for otherwise debating the DH is just a futile exercise in miscommunication. Table 2, which comprises the Appendix, summarizes the framework by listing key terms and their meanings as deployed here.

## 2. Some examples

A first task is to sketch some representative examples of dynamical cognitive science, to serve as a backdrop for the following discussion. Limitations on space dictate brevity; readers are encouraged to visit the original sources for fuller treatment.

Consider how we come to make choices between actions with various possible outcomes. If we were digital computers, we would symbolically represent to ourselves the various options and their outcomes, together with our estimates of the likelihood of those outcomes and their value to us. Reaching a decision would then be a matter of calculating the most promising option. An alternative Humean account has been proposed by psychologists Jerome Busemeyer and Jim Townsend (1993). In their Decision Field Theory (DFT) model, relevant aspects of the decision situation are represented not by symbols but by means of continuous quantities. Decision making is the interdependent evolution of these quantities over time as governed by mathematical equations (as opposed to algorithms). Decisions are made when certain thresholds are passed. The scientific question

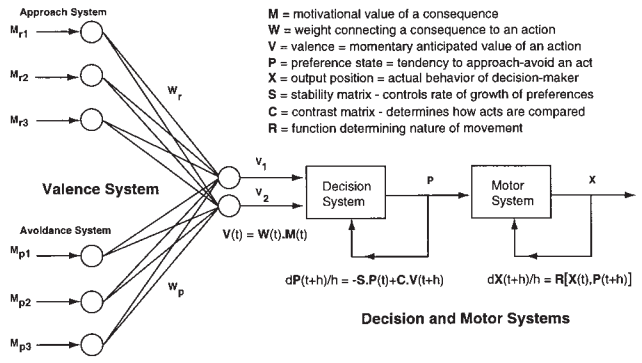



Figure 1. Outline of the Decision Field Theory dynamical model of decision-making processes (Busemeyer & Townsend 1993). The decision-making process begins (far left) with a set of possible gains and losses *(M),* filtered by attentional weights *(W)* to form the valence (momentary anticipated value, *V*) of an action. The decision system temporally integrates the valences to produce a preference state for each action *(P).* The preferences drive a motor system producing an observed action. (Figure and legend adapted from Busemeyer & Townsend 1995.)

then is: Which kind of model best accounts for the actual psychological data on human decision making? Busemeyer and Townsend claim that their model predicts actual decisions better than any "static-deterministic" model, as well as describing temporal properties of decision processes beyond the scope of traditional models.

For an example of a very different kind, consider how we manage to move our limbs. A Hobbesian would maintain that we calculate how and when to contract muscles much as a digital computer lands a 747 by calculating engine thrust, flap angle, and the like. A dynamical alternative has been under development by Scott Kelso and coworkers. His classic example is coordinating the wagging of your index fingers. Performance on this task has some remarkable properties. At low wagging speeds there are two comfortable coordination patterns, inphase and antiphase (bistability). As speed is gradually increased, antiphase patterns start to lose their stability; eventually a point comes where only inphase patterns are stable (bifurcation). As speed decreases, antiphase patterns become possible again, but not until somewhat below the original collapse point (hysteresis). Kelso found that these and other properties can be described and predicted in detail by assuming that a single, continuous, high level, "collective" variable – relative phase – evolves in a way governed by a suitable form of a simple differential equation (Kelso 1995, p. 55).[2] Variants of this "HKB"(Haken-Kelso-Bunz) model have been applied in diverse cognitive domains.[3] The basic insight is that coordination is best thought of not as masterminded by a digital computer sending symbolic instructions at just the right time, but as an emergent property of a nonlinear dynamical system self-organizing around instabilities.

These models purport to provide the best available empirical accounts of phenomena in their domains. Whether they succeed is an interesting question for specialists to address. What matters here is that they nicely illustrate the dynamical approach to cognition.

## 3. Systems, digital computers, and dynamical systems

A critical step in articulating the DH is stating, in a reasonably precise yet flexible way, just what dynamical systems are. A useful approach is to distinguish dynamical systems and digital computers as different kinds of *systems.*

### 3.1. Systems

Systems are here taken to be sets of interdependent variables.[4] A *variable* is simply some entity that can change, that is, be in different states at different times. Variables are *interdependent* when the way any one of them changes depends on the others, and change in the others depends on it. The *state* of the system is simply the state or value of all its variables at a time; the *behavior* of the system consists of transitions between states.

For example, the solar system of classical mechanics is the set of positions and momentums of the sun and planets; these are the quantities whose behaviors are described by Newton's laws. Note that the variables of the solar system in this sense are *properties* of the sun and planets. We must therefore distinguish objects (parts of the world such as the sun and planets, Macintoshes, and cognitive agents) from


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

the systems they *instantiate*. Any given object will usually instantiate a great many systems of different kinds.

*Concrete* systems are those, like the solar system, whose variables are actual features of the real world changing in real time in accordance with natural laws. *Abstract* systems are just sets of abstract variables governed by mathematical rules. Concrete systems can *realize* abstract systems. For example, two HP61 calculators realize exactly the same abstract computational system.

Concrete systems are slices of the causal organization of nature. Causal organization comes in many kinds and at many levels. Distinct systems can be intimately related. Compare the classical solar system with the system made up of all the positions and momentums of all the constituent subatomic particles. The (macro)variables of the former are built up out of the (micro)variables of the latter. The relationship between these systems is neither identity[5] nor instantiation. In what follows, a lower level system will be said to *implement* a higher level system when the variables of the latter are somehow constructed out of variables of the former. Note that implementation licenses us to identify the behavior of the one system with the behavior of the other, despite failure of strict identity between the systems themselves.

Often, change in a system depends on factors outside the system itself (e.g., the force of gravity), referred to here as *parameters.* Sometimes, changes in a parameter depend in turn on the system itself. For example, the position of the moon both depends on and affects the position of the planets. This kind of reciprocal, direct dependence is known as *coupling.* System variables and coupled parameters can be regarded as forming a larger system. This illustrates the semiarbitrariness of systems. It is always up to us to nominate a set of concrete variables as the system we will study. Reality determines whether that set is in fact a system, and how it behaves.

All systems in the current sense change in *time.* In general, time is just some intrinsically ordered set, or *order,*[6] serving to provide orderings over other things. The real time of concrete systems is the set of instants at which things can actually happen, ordered by temporal priority (before/after). Concrete events are paired with instants or periods of time, and hence stand in temporal relations with each other. Abstract systems are not situated in real time at all, and so must take some other set as their time set; usually, it is the positive integers or the real numbers. The mathematical rule imposes orderings over states of the system by pairing them with members of this set.

### 3.2. Digital Computers

The CH has benefited from considerable philosophical scrutiny. One result is a remarkable level of consensus over its basic commitments.[7] In particular, it is widely agreed to maintain that cognitive agents are *digital* computers. But what is a digital computer, as a kind of system?

A *computer* is simply anything that computes in some way or other. *Computing* is an informal notion; the basic idea is that of a process systematically transforming "questions" into "answers" – inputs into outputs, start states into final states, and so forth. The *function* computed by that process is the set of question/answer pairs themselves, or the set of pairs of entities they represent. In this general sense pretty much anything can be construed as a computer. Computation gets interesting only when significant constraints are placed on the kinds of processes involved. In classical computation theory, the standard approach has been to require that processes be *effective,* that is, produce their results by means of a finite number of basic operations specified by an algorithm (a finite recipe, or set of instructions specifying basic operations).

Digital computers, in the sense that matters for cognitive science, are systems that carry out effective computation over representations. That is, they are systems whose behaviors are algorithmically specified finite sequences of basic operations constituting manipulations of representations. This characterization can be broken down into four fundamental requirements on a system to count as a digital computer:

*(1) Digital variables and states.* First, for each variable there must be some set of *discrete* values that the variable instantiates *digitally* for the purposes of system behavior. In the concrete case, this means that the variable must instantiate those variables positively and reliably.[8] When all variables in a system are digital, the system's states are also digital. The basic operations required by effective computation correspond to digital state transitions.

*(2) Time as discrete order.* The time set must be a discrete order whose elements are the times at which the system digitally occupies its states. In abstract systems, this is usually the positive integers. In concrete systems, it is the set of periods of real time at which the machine digitally instantiates its states, as rendered discrete by the flux of transition between states. These are indexed by the positive integers ($t_1$, $t_2$, etc.).

*(3) Algorithm.* Effective computation requires basic operations to be specified by an algorithm, that is, a finite recipe specifying state transitions solely on the basis of digital properties of states. For example, the infinite range of behaviors of a Turing machine is governed by its machine table, a finite set of instructions expressed only in terms of the digital values of tape squares, head position, and head state. In concrete systems, this rule must capture one level of causal organization. That is, the transitions described by the rule must happen the way they do *because* the states bear the digital properties in terms of which the rule is expressed.

*(4) Interpretation.* The system's states and behaviors must yield to systematic interpretation. That is, there must be some domain and correspondences between the system and the domain, such that (a) the correspondences are systematic with respect to those digital aspects of the system in terms of which the rule governs system behavior and (b) the system's states and behaviors make sense in the light of those correspondences.[9]

The distinction made above (sect. 3.1) between the solar system of classical mechanics on one hand and the sun and planets on the other is mirrored by a distinction between digital computers and the ordinary notion of computers as what you take out of the box and plug into the wall. The digital computer *system* is the object of theoretical interest. The hunk of silicon, plastic, glass, metal, and the like *instantiates* some digital computer (system), and of course many other systems as well.

### 3.3. Dynamical systems

By comparison with the CH, the DH has been starved of attention.[10] Partly as a result, there is no established con-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Table 1. *Some examples of common definitions of the term "dynamical system" from outside cognitive science, arranged roughly in order, from older narrower definitions to more recent wider ones.*

| Guiding idea | Examples |
|---|---|
| 1. A system of bodies whose motions are governed by forces. Such systems form the domain of dynamics considered as a branch of classical mechanics. | "A collection of a large number of point particles." (Desloge 1982, p. 215)<br>Webster's: "Dynamics . . . a branch of mechanics that deals with forces and their relation primarily to the motion . . . of bodies of matter." |
| 2. A physical system whose state variables include rates of change. | "In the original meaning of the term a dynamical system is a mechanical system with a finite number of degrees of freedom. The state of such a system is usually characterized by its position . . . and the rate of change of this position, while a law of motion describes the rate of change of the state of the system." (*Encyclopaedia of Mathematics* 1989, p. 328) |
| 3. A system of first-order differential equations; equivalently, a vector field on a manifold. | A dynamical system is "simply a smooth manifold $M$, together with a vector field $v$ defined on $M$." (Casti 1992, p. 109) |
| 4. Mapping on a metric space. | "A *dynamical system* is a transformation $f:\mathbf{Z}\rightarrow\mathbf{Z}$ on a metric space ($\mathbf{Z}$, $d$)." (Barnsley 1988, p. 134). |
| 5. State-determination. | "A dynamical system . . . is one whose state at any instant determines the state a short time into the future without any ambiguity." (Cohen & Stewart 1994, p. 188) |
| 6. Any mapping, equation, or rule. | "A dynamical system may be defined as a deterministic mathematical prescription for evolving the state of a system forward in time." (Ott 1993, p. 6) |
| 7. Change in time. | "A dynamical system is one which changes in time." (Hirsch 1984, p.3)<br>"The term *dynamic* refers to phenomena that produce time-changing patterns . . . the term is nearly synonymous with time-evolution or pattern of change." (Luenberger 1979, p. 1) |

sensus over what dynamical systems are for the purposes of the hypothesis. Unfortunately, there is also a wide range of definitions in mathematics and science more generally (Table 1). These range from older, narrow definitions in terms of particles governed by forces to more recent, broad definitions that subsume all systems in the current sense. There is no single official definition waiting to be lifted off the shelf. Nevertheless, cognitive scientists do have a good working grasp of the issue. In the vast majority of cases they agree whether a system counts as dynamical in the sense that matters for them. The challenge here is to articulate that intuitive understanding.

An obvious feature distinguishing dynamical models in cognitive science from standard computational models is that their variables are *numerical.* One reason numbers are so useful in science is that they have *quantitative* properties. This suggests that dynamical systems in cognitive science might be defined as quantitative systems. Roughly, a system is quantitative when there are *distances* in state or time so that these distances matter to behavior. This can be true in progressively deeper ways, giving rise to progressively more substantial senses in which a system can count as dynamical.

*(1) Quantitative in state.* First, there can be distances between any two overall states of the system such that the behavior of the system depends on these distances. More precisely, a system is quantitative in state when there is a metric[11] over the state set such that behavior is systematically related to distances as measured by that metric. Such systems will be governed by a rule compactly specifying this distance-dependent change. For example, the difference equations in the DFT model describe how the system changes by telling us the distance between the values of variables at time $t$ and their values at time $t+h$.

Standardly, the relevant quantitative properties of state sets are derived from quantitative properties of the variables. Quantitative variables can be either abstract or concrete. For example, the variable $\phi$ in the HKB model is an abstract mathematical magnitude whose values are real numbers. This variable corresponds (via measurement; see Krantz et al. 1971) to a concrete quantity whose values are relative phases of oscillation of index fingers. The model works precisely because the quantitative properties of the concrete variable are reflected in the quantitative properties of the abstract counterpart.

*(2) Quantitative state/time interdependence.* A system is quantitative in *time* when time is a quantity, that is, there is a metric over the time set such that system behavior is systematically related to distances as measured by that metric. At least in cognitive science practice, systems that are quantitative in time are also quantitative in state, and these properties are interdependent. That is, the behavior of the system is such that *amounts* of change in state are systematically related to *amounts* of elapsed time. Such systems are governed by a rule specifying a quantitative relationship between change in state, elapsed time, and current state. In concrete systems, this rule captures causal organization; that is, the system changes as it does because system variables have the quantitative properties in terms of which the rule is expressed. When both state and time are quantitative, the system exhibits *rates* of change. Systems that are interdependently quantitative in state and time are governed by rules specifying the rate of change in terms of current state (e.g., first-order differential equations).


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

*(3) Rate dependence.* Third, in some systems rates of change depend on current rates of change. In these systems, variables include both basic variables and the rates of change of those variables. The solar system is a classic example. Systems whose behavior is governed by rules most compactly expressed as sets of higher order differential equations are quantitative in this sense.

In what follows, a system is taken to be dynamical to the extent that it is quantitative in one of the above senses.[12] At least four considerations support this approach. First, it reflects the actual practice of cognitive scientists in classifying systems as dynamical or not, or as more or less dynamical. Second, it sits comfortably with existing definitions. The levels of quantitative character roughly correspond to definitions 1 – 4 of Table 1. Third, it is cast in terms of deep and theoretically significant properties of systems. For example, a system that is quantitative in state is one whose states form a *space,* in a more than merely metaphorical sense; states are *positions* in that space, and behaviors are paths or trajectories. Thus quantitative systems support a geometric perspective on system behavior, one of the hallmarks of a dynamical orientation. Other fundamental features of dynamical systems, such as stability and attractors, also depend on distances. Fourth, the definition sets up a contrast between dynamical systems and digital computers (see sect. 6). For these reasons, defining dynamical systems as quantitative systems facilitates articulation and defense of the DH.[13]

## 4. The dynamical hypothesis

What does it mean to say that cognitive agents are dynamical systems? First, note that the hypothesis has two major components. The *nature* hypothesis is a claim about the nature of cognitive agents themselves; it specifies what they *are* (i.e., dynamical systems). The *knowledge* hypothesis is a claim about cognitive science: namely, that we can and should *understand* cognition dynamically. Obviously, these are closely related; the best evidence for the former would be the truth of the latter. Nevertheless, they make different claims, and are best elaborated separately.

First, some preliminary points. The proper domain of the DH is *natural* cognitive agents – that is, evolved, biological agents such as people and other animals. It need take no stand on the possibility of artificial cognition in digital computers. Second, the DH is limited in its explanatory pretensions. It is concerned only with the causal organization of agents insofar as they exhibit cognitive performances. Other forms of explanation may also be deeply illuminating. For example, evolutionary explanations might best explain *why* an agent has a particular causal organization.

What is it to be *cognitive?* In the most traditional sense, cognitive processes are those involving *knowledge;* cognitive science would then be the study of knowledge-based processes. However, as cognitive science has matured, it has diversified. Knowledge is now only one indicator of cognitive status; others include intelligence, adaptability, and coordination with respect to remote states of affairs. The concept now resists capture in terms of any concise set of strict conditions. This paper simply takes an intuitive grasp of the issue for granted. Crudely put, the question here is not what makes something cognitive, but how cognitive agents *work.*

### 4.1. The nature hypothesis

The nature hypothesis tells us what cognitive agents are by specifying the relation they bear to dynamical systems. It is common to interpret the hypothesis as asserting that cognitive agents are literally identical with some particular low-level system made up of a large number of internal, low-level quantities such as neural firing rates. However, this needs correction in almost every respect.

First, the relationship at the heart of the nature hypothesis is not identity but *instantiation.* Cognitive agents are not themselves systems (sets of variables) but, rather, objects whose properties can form systems. Cognitive agents instantiate numerous systems at any given time. According to the nature hypothesis, the systems responsible for cognitive performances are dynamical.

Second, cognitive agents "are," in this sense, not some particular dynamical system but as many systems as are needed to produce all the different kinds of cognitive performances exhibited by the agent. Consider the DFT and HKB models from section 2. These models invoke quite different sets of variables. One model suggests that cognitive agents make decisions by virtue of change in valences, preferences, and so forth; the other, that cognitive agents coordinate finger movements by virtue of change in relative phase. These models are not in competition. Both might be complete accounts of phenomena in their respective domains, implying that cognitive agents are many dynamical systems at once.

Another noteworthy fact about these models is that the variables they posit are not low level (e.g., neural firing rates) but, rather, macroscopic quantities at roughly the level of the cognitive performance itself. The lesson here is that the nature hypothesis is concerned in the first instance not with low-level systems but with how agents are causally organized at the highest level relevant to an explanation of cognitive performances, whatever that may be.

Finally, notice that the DFT model includes not only "internal" variables such as preferences and valences, but also the "position" of the agent. More generally, the dynamical system responsible for a given kind of cognitive performance might include variables not literally contained within the agent itself, on any ordinary conception of its boundaries. For example, ecological psychologists understand visually guided locomotion as change in a dynamical system that includes aspects of both the organism and the environment (e.g., the optic flow; Warren 1995).

### 4.2. The knowledge hypothesis

It is one thing for cognitive agents to *be* dynamical systems, but it is quite another for us to *understand* them as such. The knowledge hypothesis is the bold claim that cognitive science *can* and *should* take dynamical form. What does this involve?

**4.2.1. Dynamical models.** Given something we wish to understand – an explanatory *target* – a model is some *other* thing, relevantly similar but somehow more amenable to investigation. Understanding of the model transfers to the target across the bridge of similarity. Note that often the full complexity and detail of the target will defy human comprehension. In such cases, a model provides scientific insight precisely because it is a simplification.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

One of the most common strategies in science is the use of abstract dynamical systems as models. The dynamical approach to cognition follows in this tradition. The performance of interest is taken to be interdependent change in some concrete dynamical system instantiated by the agent. The scientist furnishes an abstract dynamical system to serve as a model by specifying abstract variables and governing equations. Simple models can be fully understood by means of purely mathematical techniques. More commonly, however, scientists enlist the aid of digital computers to *simulate* the model (i.e., compute approximate descriptions of its behavior). The simulation results are compared to experimental data from the target. To the extent that the correspondence is close, the target system is taken to be similar in structure to the abstract dynamical model. Note that the digital computer, because it is not itself a dynamical system (for explanation of this claim, see sect. 6.2), is not similar in the relevant sense to the target system, and so is not a model of it. We do not attempt to understand the target by understanding the *digital computer*; rather, we use the computer as a tool in our attempt to understand the target by understanding the *abstract model*.

The distinctive flavor of Humean dynamical modeling is enhanced by juxtaposition with its Hobbesian counterpart (Fig. 2). In both cases, there is a target system, an abstract model, and a digital computer. In the latter case, however, the target is assumed to be a digital computer; the abstract model is not a dynamical system but a digital computer; and the concrete digital computer does not *simulate* but rather *realizes* the abstract system. Indeed, the abstract model is often specified *by* providing the concrete computer which realizes it. Because they are identical in computational structure, both will be relevantly similar to the target if either is; therefore, both abstract and concrete systems count as models.

The basic structure of dynamical modeling is nicely illustrated by Busemeyer and Townsend's work. There are many parallels with classical mechanics. Such work comes perhaps closest to realizing the Humean dream. However, it would be misleading to suggest that dynamical modeling in cognitive science is stuck in the mold of classical physics. Obviously, cognitive phenomena differ in important ways from ordinary physical phenomena. Dynamical cognitive science has had to generate its own variations on traditional practices of dynamical modeling. Dimensions along which such variation is found include the following. (1) To what do model variables correspond? The quantities invoked in dynamical accounts often differ fundamentally from ordinary physical quantities. "Valence" and "preference," for example, do not appear in textbooks of mechanics. (2) At what level is the correspondence with the target? In physical models individual variables are usually taken to correspond directly to concrete physical quantities. In dynamical modeling in cognitive science, there might be *no* concrete quantity corresponding to individual variables. The correspondence between model and reality is at higher levels of dynamical structure. Individual units of a connectionist model, for example, may be significant only insofar as they support attractors that do correspond to aspects of cognition, such as a recognition state. (3) Is the correspondence quantitative or qualitative? Physical models are generally expected to match empirical data in more or less precise quantitative detail. A model of global warming, for example, should tell us exactly how much average temperature will rise. Such virtue is less common in dynamical cognitive science: as often as not, models match data qualitatively, at some level of abstraction. (In this respect dynamical modeling apes computational modeling.)

**4.2.2. Dynamical tools.** Understanding cognitive agents as dynamical systems means more than just using certain kinds of models. Those models, and also the cognitive performances themselves, must be understood *dynamically*. Roughly, this means taking the resources of *dynamics* – as opposed, for example, to mainstream computer science – as the basic descriptive and explanatory framework. But what are those resources?

Within dynamics there is a convenient distinction between dynamical modeling, on one hand, and dynamical systems theory (DST) on the other. Dynamical modeling is

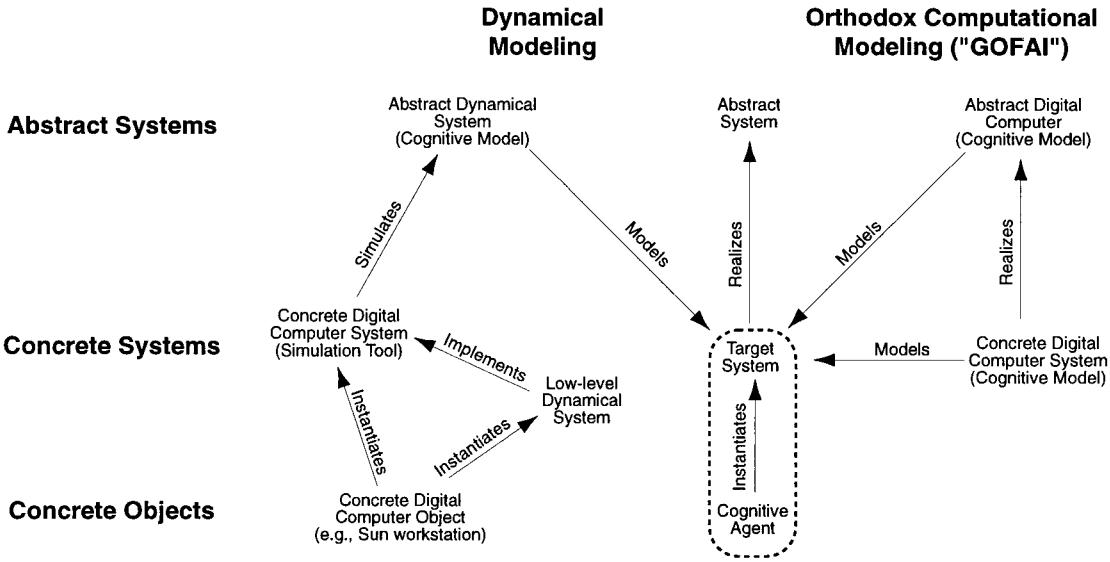



Figure 2. Basic structure of dynamical modeling as opposed to the kind of computational modeling found in mainstream cognitive science.


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

a branch of applied mathematics; its concern is to understand natural phenomena by providing abstract dynamical models. The skeletal structure of such modeling was described in the previous section. The theory of dynamical modeling is a powerful repertoire of concepts, proofs, methods, and so forth, for use in this activity. DST, on the other hand, is a branch of pure mathematics. Its domain extends to any kind of describable change, but it focuses attention particularly on systems for which there is no known way to specify behaviors as functions of time (e.g., systems whose rule is a set of nonlinear differential equations with no solutions). The fundamental move is to conceptualize systems *geometrically,* that is, in terms of positions, distances, regions, and paths in a space of possible states. DST aims to understand structural properties of the *flow,* that is, the entire range of possible paths.[14]

There is no clear line between these two sides of dynamics, but the contrast is significant. Hume envisioned psychology as dynamical modeling, but that alone does not suffice. The distinctive complexities of cognition yield to scientific understanding only when dynamical modeling is enriched by the perspective and resources of DST. Poincaré pioneered DST late last century, but the bulk of it has only been developed in the last few decades. Contemporary dynamics would be a whole new subject to Newton or even Maxwell. Hume aspired to be the Newton of the mind, but in hindsight Poincaré would have made a better model.

Dynamics plays much the same role in dynamical cognitive science as computer science (the theory of computational systems, particularly digital computers) plays in traditional cognitive science. Computer science is not itself a theory of cognitive processes. Rather, it provides a powerful set of tools for use in developing accounts of particular aspects of cognition. Therein lies the hard empirical work of mainstream cognitive science. Likewise, dynamics does not somehow automatically constitute an account of cognition. It is a highly general framework that must be adapted, supplemented, fine-tuned, and so forth, to apply to any particular cognitive phenomenon. This typically involves merging dynamics with other constructs (e.g., the schema; Rumelhart et al. 1986b) or theoretical frameworks (e.g., ecological psychology; Turvey & Carello 1995). Some authors have argued for even more dramatic reorientations in our understanding of dynamical systems for the purposes of understanding biological or cognitive systems. See, for example, the work of Robert Rosen on "anticipatory systems" (Rosen 1985) and George Kampis on "component systems" (1991).

Contemporary dynamics provides powerful resources for describing general properties of the behavior of systems. These resources can be brought to bear even in the absence of an actual equation-governed model. If done rigorously, this can buy a qualitative or preliminary understanding of the phenomenon, which may be the best available and forms a solid foundation for further exploration.[15] This approach is useful in situations where, for whatever reason, providing a model is not currently feasible (e.g., Thelen 1995).

**4.2.3. Dynamical perspective.** At the highest level, there are a number of general characteristics of a broadly dynamical perspective on some natural phenomenon. The following stand out particularly strongly when the subject is cognition and the contrast is with a computational approach:

**4.2.3.1. Change versus state.** Change and state are like two sides of one coin. Nevertheless, theoretical perspectives can differ in their primary emphasis or focus. Dynamicists are interested, in the first instance, in how things change; states are the medium of change, and have little intrinsic interest. Computationalists, by contrast, focus primarily on states; change is just what takes you from one state to another.

**4.2.3.2. Geometry versus structure.** How are states of a system conceptualised? Computationalists focus on internal structure, and in particular on internal combinatorial or syntactic structure – how basic pieces are combined to form structured wholes. Dynamicists, by contrast, understand a state geometrically, in terms of its *position* with respect to other states and features of the system's dynamical landscape such as basins of attraction. In other words, they focus on where the state is, rather than what it is made of.

**4.2.3.3. Structure in time.** Sophisticated cognition demands structural complexity in the cognitive system. How is that structure realized? Computationalists tend to think of it as laid out statically – as all present at one time – and of cognition as simple transformations of static structures. DST suggests an alternative. Systems with simple states – perhaps just one numerical variable – can behave in very complex ways. This enables dynamicists to think of cognitive structure as laid out temporally, much like speech as opposed to the written word. Cognition is then seen as the simultaneous, mutually influencing unfolding of complex temporal structures.

**4.2.3.4. Timing versus order.** Dynamicists tend to be interested in how behaviors happen in time, whereas computationalists are interested in what the behavior is, regardless of timing details. Put another way, computationalists focus on *which* states the system passes through, whereas dynamicists focus relatively more on *when* it passes through them.

**4.2.3.5. Parallel versus serial.** Dynamicists tend to think of systems as operating in parallel, that is, all aspects changing interdependently at the same time. Computationalists, by contrast, tend to think of systems as serial: most variables remain unchanged in any given state transition. For a dynamicist, change is standardly global; for a computationalist, change is standardly local.

**4.2.3.6. Ongoing versus input/output.** Computationalists standardly think of a process as commencing with an input to the system. The task for the system is to produce an appropriate output, and it does so via a sequence of internal operations culminating in the system's halting with that output. Dynamicists, by contrast, think of processes as always ongoing, not starting anywhere and not finishing anywhere. The goal is not to map an input at one time to an output at some later time, but to constantly maintain appropriate change.

**4.2.3.7. Interaction.** State-setting or coupling? How does a cognitive system interact with other things, such as the environment? Computationalists standardly think of interaction as setting *state;* the system changes in its own way from that state, until new input resets state again. Dynamicists recognize an alternative: interaction can be a matter of pa-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

rameters influencing the shape of change. Input is conceived of as an ongoing influence on the direction of change, and output as an ongoing influence on something else, just as a radio set is continuously modified by an incoming signal and at the same time is delivering its sound. Sometimes interaction is a matter of coupling – two systems simultaneously shaping each other's change.

**4.2.3.8. Representations.** Standard explanations of how systems come to exhibit sophisticated cognitive performances advert to internal representations. Computationalists take representations to be static configurations of symbol tokens. Dynamicists conceive representations very differently. They find their representations among the kinds of entities that figure in DST, including parameter settings, system states, attractors, trajectories, or even aspects of bifurcation structures (e.g., Petitot 1985a). Currently, most dynamicists make use of only the tip of the theoretical iceberg that is dynamics. As dynamical modeling increases in mathematical sophistication, we can expect representations to take even more exotic forms.

**4.2.3.9. Antirepresentationalism.** Unlike digital computers, dynamical systems are not inherently representational. A small but influential contingent of dynamicists have found the notion of representation to be dispensable or even a hindrance for their particular purposes. Dynamics forms a powerful framework for developing models of cognition that sidestep representation altogether. The assumption that cognition must involve representations is based in part on inability to imagine how any nonrepresentational system could possibly exhibit cognitive performances. Within the dynamical approach, such systems can not only be imagined, they can be modeled and constructed (see, e.g., Beer 1995a; 1995b; Freeman & Skarda 1990; Harvey 1992; Husbands et al. 1995; Skarda & Freeman 1987; Wheeler 1994).

### *4.3. The dynamical hypothesis, exposed*

Summarizing these points yields the following compact formulation of the DH: for every kind of cognitive performance exhibited by a natural cognitive agent, there is some quantitative system instantiated by the agent at the highest relevant level of causal organization, so that performances of that kind are behaviors of that system; in addition, causal organization can and should be understood by producing dynamical models, using the theoretical resources of dynamics, and adopting a broadly dynamical perspective.

## 5. Considerations favoring the dynamical hypothesis

What can be said in favor of the DH? Specific aspects of cognition generate idiosyncratic cases for dynamical treatment, but our interest here is in general considerations. Space limits preclude complete coverage, but the following arguments are among the most important.[16]

Most obviously, there is a kind of empirical success argument, paralleling Newell and Simon's primary argument for the CH.[17] It starts from the impressive track record of dynamics itself. Dynamics is arguably the most widely used and powerful explanatory framework in science. An extraordinary range of natural phenomena have turned out to be best described as – that is, to be – a matter of interdependent coevolution of quantitative variables. It would hardly be surprising if dynamics found application in the study of cognition as well. Michael Turvey for one has long been arguing that the proper road to a deep understanding of natural cognition is to strive patiently to extend and apply the tried and true techniques of natural science to incrementally more complex biological and cognitive phenomena (see, e.g., Swenson & Turvey 1991; Turvey & Carello 1981).

The empirical success argument in the form just presented has little weight on its own, for cognition differs from other phenomena in important ways. Its force really comes into play when combined with evidence of success in cognitive science itself. There is now a considerable amount of such evidence, some of which has already been cited. Of course, the claim is not that there is *now* sufficient empirical evidence to establish the supremacy of the DH. Indeed, there are numerous aspects of cognition for which, considered in isolation, the case for dynamical treatment is currently weak at best. The argument is that such successes as do exist, in conjunction with the general track record of dynamics, augurs well for the DH. The two levels of the argument require and reinforce each other.

What *explains* any success the dynamical approach has exhibited thus far? And what underpins confidence that there will be more? The foremost consideration is simply *that natural cognition happens in real time.* This blunt fact is multifaceted. Every cognitive process unfolds in continuous time, and the fine temporal detail calls out for scientific accounting. Moreover, many cognitive structures are *essentially* temporal: like utterances, they exist only as change in time. Often, getting the timing right is critical to the success of cognitive performance; this is especially so when in direct interaction with surrounding events.

Hobbesian computational models have made a bet that cognitive phenomena can be described in a way that abstracts away from the full richness of real time, replacing it with discrete orderings over formal states. From a dynamical perspective, this looks ill advised. Dynamics, by contrast, takes the nature of change in time as its primary focus. It is the preeminent mathematical framework for description of temporal phenomena. Taking cognitive agents to be dynamical systems allows scientific explanation to tap into this power.

A third argument focuses on the *embeddedness* of cognition. Even the loftiest forms of natural cognition are in fact embedded three times over: in a nervous system, in a body, and in an environment. Any account of cognition must eventually explain how it is that cognition relates to that which grounds and surrounds it. Now, suppose the behavior of brain, body, and environment all turn out to be best described in dynamical terms. Suppose, in short, that cognition is thoroughly embedded in dynamics. The challenge would then be to explain how cognitive phenomena are constituted of, shaped by, and interact with those dynamical phenomena. Explaining embeddedness is never trivial, and it stands to reason there will be greater problems in relating systems of fundamentally different kinds than in relating systems of fundamentally the same kind. Mainstream computational cognitive science has for the most part simply shelved problems of embeddedness, preferring to study cognition independently of its neurobiological realization, and treating the body and environment as belonging on the far side of occasional symbolic inputs and outputs. When


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

embeddedness is confronted head-on, dynamical accounts of cognition immediately become attractive. For example, one virtue of the Catherine Browman and Louis Goldstein dynamical *phonology* (1992) is that it integrates directly with Elliot Saltzman's dynamical model of *speech coordination* (Saltzman & Munhall 1989). Dynamical cognition sits comfortably in a dynamical world.

A fourth argument focuses on the *emergence* and *stability* of cognition. Investigation of some complex phenomenon can always take at least two directions: What is it like? And, how does it get – and stay – that way? In the case of cognitive mechanisms and processes, we can address their nature, or how it is they arise and are sustained. In the long run, our answers to these questions must hang together. Natural cognitive agents exhibit extraordinary levels of structural complexity, yet there are no architects or engineers responsible for building and maintaining that structure. The generic name for the answer to the problem of the emergence and stability of cognition is *self-organization.* Self-organization of interesting kinds of complex order appears to require systems in which there is simultaneous, mutually constraining interaction between large numbers of components. DST is the dominant mathematical framework for describing the behavior of such systems. In short, the claim is that we must understand cognitive agents as dynamical systems, because only in that way will our account of what cognition is be properly integrated with our account of how the world sustains any of it.

Each of these lines of thought was cast in the form of an attempt to demonstrate that the DH is basically *true.* With respect to that goal, they are obviously not "knock down" arguments. They do, however, indicate that the hypothesis is worthy of sustained empirical investigation of precisely the kind that has been and is being conducted, and that forms the basis of the formulation of the DH presented here.

## 6. The general objections

This section considers a selection of general objections to the DH as an open empirical hypothesis. As John Stuart Mill said, "three-fourths of the arguments for every disputed opinion consist in dispelling the appearances which favour some opinion different from it."[18] Addressing these objections is also a useful way to elaborate and clarify the hypothesis.

The objections considered fall into two main categories: those purporting to show that the DH is not a genuine *alternative* to the CH, and those purporting to show that it is not open, that is, its empirical inadequacy is somehow already determined. All amount to sweeping attempts to dismiss or downplay the DH in advance of detailed empirical investigation. All mix insight with confusion to produce plausible but misguided attacks.

### 6.1. The "trivially true" objection

> Everything is a dynamical system. Cognitive agents must be dynamical systems at some level. The DH is trivially true, and makes no substantial claim about the nature of cognition.

This objection is mostly bluff. No doubt there is *some* vague sense in which it could be said that everything is a dynamical system. Properly interpreted, however, the DH makes a much more specific claim.

On the one hand, according to the nature hypothesis, cognitive agents instantiate quantitative systems at the highest relevant level of causal organization. It may be trivial that every cognitive agent instantiates some dynamical system or other. It is certainly not trivial that every cognitive performance is at the highest level a dynamical phenomenon. This is not true of ordinary digital computers, and according to the orthodox CH, it is not true of people.

On the other hand, according to the knowledge hypothesis, cognition can be *understood* in dynamical terms. If *this* were trivially true, cognitive science would have been completed long ago. In practice, it is very challenging to establish that some aspect of cognition can be understood dynamically. Patient steps in this direction are the stuff of which whole careers are made. Some of the greatest achievements in science have amounted to describing some natural phenomenon (e.g., celestial motion) in dynamical terms. This activity is no more trivial in cognitive science than anywhere else.

### 6.2. The "false opposition" objection 1: Computers are dynamical systems

> Ordinary electronic computers are dynamical systems. In general, digital computers are dynamical systems as well. The DH is therefore not an interesting alternative to the CH.

This objection gains plausibility by mixing together at least three distinct lines of thought. Each is based on a different reason for thinking that digital computers are dynamical systems. Each has elements of truth but also problems.

**6.2.1. Digital computers are state determined, rule governed.** A first line of thought takes digital computers to count as dynamical systems because they satisfy some broad definition; for example, they are state-determined systems, or they are governed by some mapping, and the like. This kind of move is reasonable in the light of some strands of contemporary usage (see Table 1). However, it only appears to constitute an objection to the DH because it equivocates on the term "dynamical system." The DH takes cognitive agents to be dynamical systems in a much more specific sense, that is, quantitative systems.

**6.2.2. Digital computers are quantitative systems.** A second line of thought does not equivocate. Rather, it suggests that the definition of dynamical systems as quantitative systems is broad enough to embrace digital computers as such.

Digital computers and dynamical systems are two classes of systems picked out by reference to different properties: roughly, effectiveness and interpretation as opposed to quantitativeness. Generally, systems exhibiting the one property fail to exhibit the other and vice versa. In a typical Turing machine, for example, there is no systematic relationship between system behavior and distances between states. A tape square's values are *different* but not relevantly *distant* from each other. System behavior turns only on which values happen to obtain (i.e., type identity), not on how far those values are from any others. Similarly in the case of time. Turing machine states are indexed by means of the positive integers. There are distances between integers, but these distances generally bear no systematic relationship to system behavior. The integers might just as well be replaced by any other sufficiently large merely ordered set, such as names in the New York City telephone directory.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Since there are generally no relevant distances in state or time in digital computers, it makes no sense to describe their behavior in terms of rates of change (not to mention dependence on rates of change). This is why in practice computer scientists don't bother with distances between states, rates of change, and so forth.

There is a common temptation to suppose that digital computers count as quantitative systems arising from the correct observation that certain metrics apply to any set of values, regardless of the nature of those values (e.g., Padulo & Arbib 1974, pp. 91–92). Thus *every* variable is a quantity, and so even digital computers have metric spaces as state sets. The crucial point, however, is that the distances measured by these trivial metrics bear no systematic relationship to system behavior. Turing machines bounce around their state spaces in ways that will seem utterly erratic until one realizes that their order is based on formal properties, not quantitative properties.

Oranges come in many kinds. Some are Valencia, some are expensive; occasionally, an orange is both. Similarly with digital computers and dynamical systems. In coincidental, contrived, or trivial cases, one and the same set of variables might satisfy the conditions for both classes. Nevertheless, digital computers and dynamical systems are classes of systems picked out by reference to fundamentally different properties. In general, systems exhibiting one property fail to exhibit the other.

**6.2.3. Digital computers are dynamical systems at the hardware level.** A third line of thought is based on the idea that all *concrete* digital computers are in fact dynamical systems at some lower level of description. For example, standard general purpose digital computers such as Macintoshes are dynamical systems at the level of electronic circuits. Now, there is truth in this, but not enough to vitiate the relevant contrast. The fundamental problem here is that "are" is too crude; it rides roughshod over a number of issues.

To sort out the relationship between digital computers and lower level dynamical systems, we must distinguish at least three different relationships: instantiation, identity, and implementation. At any given time a Macintosh *instantiates* a great many different systems at different levels. One of these is the high-level digital computer by virtue of which, for example, it calculates my taxes. Presumably it also instantiates some hugely complex electrical dynamical system. The Macintosh is not *identical* with either of these systems. Neither are they strictly identical with each other; most obviously, they have different numbers of variables. Of course, the macrovariables of the high-level digital computer are ultimately built up out of the microvariables of the electronic system, and so there is presumably some lower level dynamical system *implementing* the high-level digital computer. Thus, while there is one clear sense in which the digital computer "is" some lower level dynamical system, there is also a clear sense in which it "is" not that system.

### 6.3. The "false opposition" objection 2: Dynamical systems are computers

> Much recent research in computation theory has been exploring the computational power of dynamical systems. There is no inherent conflict between dynamics and computation, and so there is no real opposition between the computational and dynamical hypotheses.

It is true that there is no inherent conflict between dynamics and computation, but the conclusion does not follow. Again, the issues must be teased out more carefully.

Recall from section 3.2 that effective computation is a specific kind of computation, resulting from a certain kind of constraint on the processes involved. Other kinds of computation result from adopting different constraints. In particular, we can focus attention on some class of dynamical systems (Blum et al. 1997; 1989; Moore 1991; 1996). As long as there is some way to specify the "questions" and "answers," we can see dynamical processes as computing functions. For example, Hava Siegelmann has extensively studied the computational properties of one class of dynamical systems, recurrent neural networks (Siegelmann & Sontag 1994). Indeed, it can be proved that certain classes of dynamical systems are more powerful – can compute a wider class of functions – than Turing machines.[19] So, dynamical systems can compute, that is, be computers, without needing to be digital computers. This is why research into the power of dynamical systems is an interesting new branch of computation theory.

The most famous and influential of all critiques of the mainstream computational approach to cognition is surely *What Computers Still Can't Do* (Dreyfus 1992). In that book, Dreyfus noted that brains might well turn out to be "analogue" rather than digital computers. Similarly, as Churchland and Sejnowski have argued at length, biological neural networks can be understood as computing in ways that differ fundamentally from ordinary digital computation (1992). Like these perspectives, the DH can embrace the idea that cognitive processes are computational, while preserving a contrast with the CH. This does not diminish but, rather, fortifies the DH, by allowing it to incorporate computational ideas without inheriting orthodoxy's excess baggage.

### 6.4. The "false opposition" objection 3: Dynamical systems are computable

> There is no good reason to think that any cognitive process is not effectively computable. Even if cognitive agents are dynamical systems, they will still be computable systems. Therefore, it is misguided to present the DH as an alternative to the CH.

One particularly troublesome mistake is blurring the distinction between *computational* and *computable.* Just as employers and employees stand at opposite ends of an employment contract, so *computational* and *computable* stand at opposite ends of the relation *computes.* The former applies to whatever *does* the computing; the latter to whatever *gets* computed. In classical theory, a digital computer does the computing, and a function over the integers gets computed. The effectively computable functions over the integers are all and only the partial recursive functions.

Computation theorists, including Turing himself, quickly turned to asking what *else* might be effectively computed. Via arbitrarily good approximation, the purview of effective computation was gradually extended to embrace real numbers, functions over real numbers, differential equations, and so on (Earman 1986; Grzegorczyk 1957; Turing 1936). In this way, issues of effective computability can be raised for all the standard mathematical constructs of analysis and physics. Just what is and is not effectively computable rapidly becomes a rather complicated business (see, e.g., Pour-El & Richards 1989).


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Now, we can regard a *system* as computable just in case its behavior is governed by some computable function. The solar system of classical mechanics is effectively computable in this sense. Currently, as far as we can now see, most if not all dynamical systems of practical relevance to cognitive science are effectively computable.[20] This does not make those systems digital computers. Digital computers can compute functions governing systems that are not themselves digital computers. Thus, the computability of dynamical models does not destroy the contrast between the dynamical and computational hypotheses.

### 6.5. The "straw man" objection

> Turing machines are caricatures of computers. The DH is being matched against a straw man. It is not a substantial alternative to the CH as properly understood.

There are two issues here. One is whether the CH, as characterized here, is a straw man. Two considerations suffice to dispel this objection. First, the characterization offered here is just the standard philosophical account, as developed in numerous places.[21] Second, a great many models in cognitive science (e.g., those developed within the SOAR [Newell 1991] framework) do in fact conform to that account.

The other issue is whether the standard account misunderstands the "true" CH, that is, deeply misconceives computers and computational modeling in cognitive science. This may be; Brian Smith, for one, has begun formulating a critique of received wisdom in this area (Smith 1996; Smith, in preparation). These issues go beyond the scope of the present discussion. If and when some superior understanding of the CH clearly supplants the orthodox account, the relationship between the dynamical and computational hypotheses will need to be reconsidered.

### 6.6. The "description, not explanation" objection

> Dynamical models are at best descriptions of the data, and do not explain why the data take the form they do. For genuine explanation, we need computational models describing the underlying causal mechanisms.

Dynamical theories of cognitive processes are deeply akin to dynamical accounts of other natural phenomena such as celestial motion. Those theories constitute paradigm examples of scientific explanation. Consequently, there is no reason to regard dynamical accounts of cognition as somehow explanatorily defective.

Dynamical explanations typically proceed by providing equations defining an abstract model. Many factors are relevant to the goodness of a dynamical explanation, but the account should at least capture succinctly the relations of dependency, and make testable predictions. A *poor* dynamical account may amount to little more than ad hoc "curve fitting," and would indeed count as mere description. Its problem, however, is that it is poor, not that it is dynamical.

Traditional computational cognitive science offers explanations of a quite distinctive kind (Haugeland 1978), and many cognitive scientists have become so accustomed to such explanations that anything else seems inadequate. The explanations offered in dynamical cognitive science are indeed quite different (Garson 1996; van Gelder 1991), but are not for that reason inferior.

### 6.7. The "not as cognitive" objection

> Dynamics is a general purpose framework that applies to any behavior of an agent, regardless of whether that behavior is cognitive or not. Dynamics does not focus on the specifically *cognitive* aspects of systems; it does not explain cognitive performances "as cognitive." Genuine explanation in cognitive science must be framed in terms of aspects of cognitive agents other than their purely dynamical properties.

This objection concedes that dynamical explanations are nontrivial empirical explanations, and that they really are quite different from computational explanations. It challenges the nature of the explanation being offered. Dynamics is held to be too general, failing to explain cognition in terms of its distinctive features.

Underlying this objection is an important misconception about the DH. That hypothesis asserts that cognitive agents are dynamical systems of quite special kinds. Therefore, as emphasized in section 4.2.2, understanding cognitive agents *as* dynamical systems is not simply the routine application of generic dynamics to systems that happen to be exhibiting cognitive performances. It requires that the resources of dynamics be developed and supplemented in order to provide explanations of those special kinds of behaviors. Thus, dynamical cognitive science always incorporates considerations distinctive to particular kinds of cognition into dynamical frameworks to produce explanations that are fundamentally dynamical in form, but are nevertheless tailored to explain cognitive performances "as cognitive." To take just one example, Jean Petitot merges Ron Langacker's cognitive grammar with René Thom's morphodynamics to yield a thoroughly dynamical approach to syntax (Petitot 1995).

### 6.8. The "wrong level" objection

> There is an important role for dynamical descriptions in any complete account of the nature of a cognitive agent, but they are pitched too low to explain cognition.[22]

A common misconception about the dynamical approach is that it operates solely or primarily at "lower" or "micro" levels of description. In fact, dynamics is not intrinsically limited to *any* level or domain. In the natural sciences, dynamics finds application at all levels from quantum mechanics to cosmology. It gets its grip wherever sets of interdependently changing quantities are found. Similarly in cognitive science: dynamicists develop their explanations at the level of theoretical interest, whatever that might be (see sect. 4.1).

One significant difference between the dynamical approach and PDP-style connectionism turns on this point. They agree that cognitive performances are behaviors of dynamical systems. The PDP approach, however, takes those systems to be high-dimensional neural networks operating at a level below that of orthodox descriptions (Smolensky 1988); as expressed in the titles of the famous volumes,[23] they constitute the *microstructure* of cognition. The dynamical approach is more catholic; it embraces dynamical models of all kinds and at all levels.

### 6.9. The structure objection

> Sophisticated cognitive performances require complex internal structures. The dynamical approach is taking a huge step backward in trying to replace symbolic representations with quantities. To explain high-level cognition, dynamical systems will have to implement computational mechanisms.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Almost everyone now agrees that most kinds of cognitive performance can be explained only by reference to complex structures internal to the system responsible for those performances. Still, it remains an open question what form those structures might take. Hobbesian cognitive scientists are banking on the idea that they are the kind of structures found in digital computers, that is, symbol structures (Newell & Simon 1976) or "classical" combinatorial representations (Fodor & Pylyshyn 1988). Lying behind this idea is an assumption that the kinds of complex structures required cannot exist in any system except by instantiating digital symbol structures.

However, as dynamical cognitive science has matured, it has become apparent that dynamical systems can incorporate combinatorial structures in various ways without merely implementing their digital cousins (van Gelder 1990). For example, arbitrarily many structures can be mapped to states of a dynamical system such that these states can then be used as the basis of systematic processing (e.g., Chrisman 1991; Pollack 1990). Other work has found combinatorial structure in the attractor basins of appropriate dynamical systems (Noelle & Cottrell 1996), or in the trajectories induced by sequences of bifurcations ("attractor chaining" [van Gelder & Port 1994]). The possibilities have really only begun to be explored. The dynamical approach is not vainly attempting to do *without* complex internal structures. Rather, it is in the process of dramatically reconceiving how they might be instantiated.

#### 6.10. The complexity objection

> Natural languages are only effectively described by some form of context-sensitive grammar. In the standard Chomskyan hierarchy, languages of this complexity can be handled only by computers at least as powerful as linear-bounded automata (LBAs). Therefore, natural language speakers must be computers at least as powerful as LBAs.

The conclusion of this argument is ambiguous, between computers in general and digital computers. On the former interpretation, the argument is sound but fails to conflict with the DH. It was pointed out above that dynamical systems can compute, that is, be computers. The complexity of natural language constrains speakers' computational *power* but not the *kind* of computer they instantiate. It remains an open empirical question whether the computers in question are best thought of as digital or dynamical (Elman 1995).

In the latter interpretation, the argument simply equivocates. The premises establish that speakers must be computers in *some* sense; the conclusion claims they must be digital computers. The dominance of digital computers in the theory of computation, cognitive science, and computer technology has created an unfortunate tendency to confuse computers in general with digital computers. This is what drives the objection.

#### 6.11. The "Not cybernetics again!" objection

> The dynamical approach is just cybernetics returning from the dead.

What was cybernetics? Wiener famously defined it as "the science of communication and control in man and machine," but it soon developed into an even wider enterprise: a kind of general, nonreductionistic study of *systems,* particularly self-sustaining systems in their environments (see, e.g., Parsegian 1973). Throughout its brief ascendancy, cybernetics enthusiastically embraced anything of conceivable relevance to complex systems, including information theory, communication theory, automata theory, neurophysiology, systems theory, game theory, and control theory.

Dynamics was certainly mixed up in all this, and the DH is sometimes traced back to a leading cyberneticist, H. Ross Ashby. Still, the demise of cybernetics implies little about the contemporary dynamical approach, for they differ in important ways. The DH is, by comparison, tightly circumscribed. It is concerned with cognition specifically, rather than systems generally, and is defined in terms of a core commitment to a single framework. The fate of cybernetics as a whole no more attaches to the dynamical approach than it does to other disciplines with ancestral links to cybernetics, such as computational neuroscience and artificial intelligence. Moreover, much more powerful tools are available today. The bulk of DST has been developed in the period *since* cybernetics. Also, dynamicists now have on their desks computer simulation tools (hardware and software) beyond the dreams of cyberneticists. Where cyberneticists could only speculate, dynamicists can now furnish and understand complex models.

#### 6.12. The "humans compute" objection

> Humans can do arithmetic in their heads. At least some cognitive activity is specifically *digital* computation. Therefore, the DH cannot be the whole truth about cognition.

If it is granted that mental arithmetic and like processes are, literally, digital symbol manipulation inside the head, then the DH should indeed graciously concede. The *general* truth of the DH is compatible with certain special activities counting as exceptions. However, we should be wary of granting, in advance, that mental arithmetic *is* symbol manipulation. Certainly, it *seems* like symbol manipulation: numerals, lines, and so forth are "seen in the mind's eye." It does not follow that there are symbols in the head, that the states and processes subserving such "seeing" actually instantiate symbols and their manipulations. Imagining the Eiffel tower does not entail that one has the Eiffel tower, or even a picture of it, inside one's head (Ryle 1984, Ch. 8). We must not confuse the content of experience with the mechanisms implementing it. As usual, the question turns out to be the empirical one: In the long run, what kind of models provide the best account of the mechanisms underlying the relevant kind of cognitive performance?

## 7. Conclusion

The contemporary dynamical approach to cognition is part of a much wider scientific trend. In recent decades, there have been dramatic developments in the mathematics of DST, especially the theory of nonlinear systems, complexity, and chaos. At the same time, there has been exponential growth in available computing power, and the arrival of sophisticated programs for exploring dynamical systems. The result is that dynamical theorizing has come to be applied to a wide range of natural phenomena that were pre-


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

viously either ignored entirely or regarded as beyond the scope of standard forms of scientific explanation. So with cognition. The Humean dream of a dynamics of cognition can now be seriously pursued. The explanatory umbrella that covers so much of the rest of the natural world so effectively is gradually being extended to cover cognition as well.

The DH encapsulates the core commitment of the emerging dynamical approach. This target article has attempted to say what it means, and to establish its status as an open empirical hypothesis standing as a substantial alternative to the CH. It has not attempted to demonstrate that cognitive agents *are in fact* dynamical systems. There is mounting evidence that certain aspects of cognition are best thought of dynamically, but many others remain completely unaddressed. Only sustained empirical investigation will determine the extent to which the DH – as opposed to the CH, or perhaps some other hypothesis entirely – captures the truth about cognition.

## APPENDIX

Table 2. *Key terms and their meanings in the present discussion. This table has no pretensions beyond partially summarizing the particular regimentation proposed in this paper for the purpose of clarifying the DH in cognitive science*

| Term | Meaning in this paper |
|---|---|
| Variable | Anything that changes over time. |
| System | A set of variables changing interdependently. |
| Instantiation | A relation between a concrete system and some object or part of the world. An object instantiates a system when all the variables of the system are features of the object. |
| Implementation | A relation between concrete systems, obtaining when the variables of one system are somehow built up out of the variables of the other. |
| Parameter | Something outside (i.e., not a member of) a system, but on which change in the system depends. |
| Coupling | Mutual direct dependence. Variables $x$ and $y$ are coupled when the state of $x$ shapes change in $y$ and vice versa. |
| Concrete system | A system whose variables are all concrete features of the concrete world changing in real time. |
| Abstract system | A system whose variables are all abstract entities. |
| Realization | A relation between a concrete system and an abstract one, obtaining when the former has the same structure as the latter. |
| Time | Any intrinsically ordered set, serving to provide orderings over other things. Real time is the set of instants at which things can happen, ordered by priority (before/after). |
| Computer | Anything that computes (carries out computation). |
| Computation/ Computing | Transforming some kind of question (e.g., input object or start state) into some kind of answer (e.g., output object or final state). |
| Computational | Anything that computes (carries out computation). |
| Digital Computer | A computer carrying out effective computation over representations. A digital computer must have digital variables, discrete time, algorithmically governed behavior, and an interpretation. |
| Effective | Succeeding in a finite number of basic operations governed by an algorithm. |
| Computable | Capable of being computed; alternatively, being governed by a computable function. |
| Quantity | A variable with a metric over its values. |
| Dynamical System | A quantitative system. A system that is at least quantitative in state; may also be interdependently quantitative in state and time, or even rate dependent. |
| Identity | "Being the very same thing as." Identity is governed by Leibniz Law: identical things have all and only the same properties. Identity for sets – and hence for systems – is having all and only the same variables. |
| Simulate | Compute a function describing some process. |
| Dynamics | Two closely related kinds of mathematics, dynamical modeling and DST. |
| Dynamical Hypothesis (DH) | Cognitive agents are dynamical systems. See section 4.3. |
| Computational Hypothesis (CH) | Cognitive agents are digital computers. |

## ACKNOWLEDGMENTS

Significant improvements in this paper resulted from discussion with or feedback from many people. Among the most influential were John Haugeland, Robert Port, Jim Townsend, Dan Dennett, Herbert Jaeger, Tim Smithers, Robert Gregson, Clark Glymour, Brian Smith, Jeff Pressing, Marco Giunti, Scott Kelso, and BBS referees.

## NOTES

**1.** Examples: cognitive neuroscience (Amit 1989; Babloyantz & Lourenco 1994; Cohen 1992; Guckenheimer et al. 1993; Mpitsos, in press; Skarda & Freeman 1987); psychophysics (Gregson 1995); perception (Bingham et al. 1997; Grossberg & Rudd 1992; McClelland & Rumelhart 1981; Port et al. 1995); motor control (Bullock & Grossberg 1988; Saltzman 1995; Turvey 1990); developmental psychology (Smith & Thelen 1993; Thelen & Smith 1993); cognitive psychology (Busemeyer & Townsend 1993; Grossberg & Gutowski 1987; Grossberg & Stone 1986; Leven & Levine 1996; Tabor et al. 1996); situated robotics and autonomous agents research (Beer 1995b; Cliff et al. 1993; Smithers 1994a); artificial intelligence (Jaeger 1996; Pollack 1991); social psychology (Kaplowitz & Fink 1992; Vallacher & Nowak 1993); ecological psychology (Kugler et al. 1980; 1982;



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Turvey & Carello 1995); synergetics (Haken & Stadler 1990); morphodynamics (Petitot 1985b; Thom 1983; Wildgen 1982). Port and van Gelder 1995 is a representative sampling of the dynamical approach. Note that works cited here are intended as examples and pointers, rather than any kind of exhaustive or definitive listing.

**2.** The basic Haken-Kelso-Bunz equation is $\dot{\phi} = -a\sin\phi - 2b\sin 2\phi$. Here $\phi$ is the single "collective" state variable of the system; in the finger coordination model, it corresponds to the oscillation phase of one finger relative to the other. The equation specifies how relative phase changes as a function of its current value $a$ and $b$ are parameters of this system; their ratio corresponds to the rate of wagging of the fingers. The equation is such that gradual changes in $a$ and $b$ can yield just the kind of qualitative changes in relative phase found in the behavior of real subjects.

This simple "frictionless" equation is altered in various ways to generate models with better fit to experimental data. For example, fluctuations and symmetry-breaking considerations are accommodated by adding terms for noise and for differences in frequency between limbs and metronome (Kelso et al. 1990).

**3.** These include aspects of motor skill learning (Schöner et al. 1992), interpersonal coordination (Schmidt & Turvey 1994), speech perception (Tuller et al. 1994), and visual perception (Hock et al. 1993). See Kelso 1995 for an overview.

**4.** This definition accords with ordinary usage (e.g., Webster's Dictionary: "a regularly interacting or interdependent group of items forming a unified whole") and systems theory (e.g., "a set of elements standing in interrelations," [von Bertalanffy 1973, p. 55]). The stance on the metaphysical status of sets adopted here is the "set-theoretic realism" elaborated in Maddy 1995. In this account, sets of physical entities are themselves physical entities, as much part of the ordinary world as planets, people and PCs.

**5.** In set theory, set identity is a matter of having exactly the same members. A set of sets is not identical with the set of the elements of those sets. Thus, strictly speaking, a set of pairs of socks is not identical with the set of socks belonging to those pairs. Of course, there is still an obvious and important sense in which these sets are the same. In this paper, this sense is captured by the notion of *implementation.*

**6.** A nonempty set $X$ is an *order,* or is *ordered,* if there is a relation $<$ over its elements with the property that for each $x,y \in X$, either $x < y$, or $y < x$, or $x = y$.

**7.** For expressions of this consensus see, for example, Clark (1989), Copeland (1993), Dreyfus (1992), Fodor (1975), Fodor & Pylyshyn (1988), Newell (1980), Newell & Simon (1976), Pylyshyn (1984). The version of this consensus now most widely accepted as definitive is probably that laid out in Haugeland (1985). The account of digital computers here is essentially just Haugeland's definition of computers as interpreted automatic formal systems as massaged into the present framework.

**8.** See Haugelend, 1985, Chapter 2. In abstract systems, discreteness of values suffices for digitality.

**9.** What is it to "make sense"? This is a difficult issue; see Haugeland, 1985, Chapter 3, for discussion. Every digital system can be set up in systematic correspondence with some domain (such as integers and functions over them) but not all such systems have an interpretation in the current sense. The ones that do are those exhibiting a further kind of order that does or could seem patterned or reasonable *to us* (humans); thus, whether something is a digital computer is human relative.

Note that having an interpretation in the current sense may not be enough to guarantee that the system has "meaning" in some stronger sense, (and hence, perhaps, "mind"). For discussion of these issues, see Harnad (1990) and Searle (1980).

**10.** Recently, philosophers have begun to repair this neglect. See, for example, Giunti (1997), Horgan & Tienson (1996), van Gelder (1995), and van Gelder & Port (1995) for discussion more or less closely related to the current issues.

**11.** A metric over a set $X$ is a function $d: X \times X \rightarrow \mathbf{R}$ that assigns to every pair of elements $x$ and $y$ a number $d(x,y) \geq 0$ such that $d(x,y) = 0$ if $x = y$, $d(x,y) = d(y,x)$, and $d(x,y) \geq d(x,z) + d(z,y)$.

**12.** This formulation is designed to accomodate some rather special cases of dynamical systems whose behavior is generally quantitative except at certain isolated points (Gregson 1993; Zak 1990).

**13.** The concept of *dynamical system* changes over time, in cognitive science as elsewhere. Future developments might prompt broadening of the current definition. For example, cognitive scientists *may* come to use as models systems whose state sets are not metric spaces, but do possess some other kind of interesting topological structure relevant to system behavior.

**14.** For introductions to dynamical modeling, see Beltrami (1987) and Luenberger (1979). For introductions to dynamical systems theory, see Abraham and Shaw (1982) or Baker and Gollub (1990). Abraham et al. (1992), Kelso et al. (1992), and Norton (1995) are chapter-length overviews of dynamics for cognitive scientists.

**15.** If done poorly, on the other hand, it is little more than handwaving with impotent metaphors. The jargon of dynamics does, unfortunately, provide all too many opportunities for pseudoscientific masquerading.

**16.** Discussion of a wider range of considerations is found in van Gelder and Port (1995).

**17.** In their celebrated paper "Computer science as empirical enquiry," Newell and Simon argue for the computational hypothesis primarily on the basis of the success of AI in producing intelligent computers, and the success of computational cognitive science in modeling cognition. The only other argument they mention is "the absence of specific competing hypotheses." (See Newell & Simon, 1976.)

**18.** (Mill 1975, Ch. 2.) In "Computing machinery and intelligence" (Turing 1950), Turing rebuts nine objections to his stance on whether computers can think; most are not attributed to anyone in particular. This paper follows these august precedents. Except where noted, the objections are not known to have appeared in print; rather, they are based on the author's experience of reactions to the dynamical hypothesis when expounded in public presentations or in related work.

**19.** The general result that dynamical systems can have "super-Turing" capacities need not be very surprising. Digital computers are a strictly delimited class of systems, and it makes sense that classes defined by alternative sets of constraints would allow more powerful processes.

**20.** Note that *effectively computable* is a theoretical notion; it is not the same as *computable in practice.* As chaos theory reminds us, some systems will always outstrip our finite computing resources.

**21.** See note 7.

**22.** The "peripheral" objection is very similar, and is dealt with by a similar response. It maintains that dynamical explanations are concerned with peripheral aspects of cognitive agents rather than cognition itself, which is more "central."

**23.** McClelland et al. (1986) and Rumelhart et al. (1986a).



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

# Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

## Dynamicists versus computationalists: Whither mechanists?

William Bechtel

*Philosophy-Neuroscience-Psychology Program, Washington University in St. Louis, St. Louis, MO 63130;* **bechtel@twinearth.wustl.edu**
**www.artsci.wustl.edu/~wbechtel**

**Abstract:** Van Gelder's characterization of the differences between the dynamical and computational hypotheses, in terms of the contrast between change versus state and geometry versus structure, suggests that the dynamical approach is also at odds with classical mechanism. Dynamical and mechanistic approaches are in fact allies: mechanism can identify components whose properties define the variables that are related in dynamical analyses.

Van Gelder's analysis focuses on two approaches to exploring cognition – the computational and the dynamical. There is, however, a third approach – the mechanistic – which is most evident in neuropsychology and cognitive neuroscience but also in many models in cognitive psychology. The point of this commentary is to explore how the latter approach relates to the computational and dynamical perspectives discussed by van Gelder.

Van Gelder differentiates *abstract systems* and *concrete systems.* Systems are "sets of interdependent variables" (sect. 3.1, para. 1), abstract systems are mathematical objects, and concrete systems are "slices of the causal organization of nature" (sect. 3.1, para. 4). In van Gelder's terms, the relation between these is one of *instantiation:* concrete systems instantiate abstract systems. Here at first pass we can identify a difference between dynamical and mechanistic approaches: systems are the subject of dynamical analyses, whereas concrete systems are the subject of mechanistic analyses. Can more be said about the relation between the two?

As Richardson and I (Bechtel & Richardson 1993) have characterized the mechanistic approach in the life sciences, its strategy is to decompose the activities of a concrete system into component activities and to localize these in the parts of the concrete system. The emphasis on decomposition brings out one of the key components of the mechanistic approach – the move down the levels from the behavior of a concrete system to the behavior of its constituent parts. Levels are, in this view, discovered by taking concrete systems apart functionally or structurally; but what defines a level is the degree of interaction between entities – a level is a point of maximal interaction between entities (Wimsatt 1972). It is this emphasis on moving between levels that gives mechanistic approaches their reductionist character, although the sort of reduction involved is that of the scientist, not the philosopher.

Discussion of levels arises frequently in van Gelder's analysis. He emphasizes both that dynamical analyses are not restricted to the lowest level but can be developed at any level of organization, and that the dynamical account of interest to cognitive science is the one "instantiated by the agent at the highest relevant level of causal organization (sect. 4.3). But how are levels to be characterized in the dynamical perspective? Van Gelder's most detailed comments on this issue are:

> Causal organization comes in many kinds and at many levels. Distinct systems can be intimately related. Compare the classical solar system with the system made up of all the positions and momentums of all their constituent subatomic particles. The (macro)variables of the former are built up out of the (micro)variables of the latter. The relationship between these systems is neither identity nor instantiation. In what follows, a lower level system will be said to *implement* a higher-level system when the variables of the latter are somehow constructed out of the variables of the former (sect. 3.1, para. 4).

The language of being "somehow constructed" suggests that there is a difficulty here in making clear the relationship between the variables associated with entities at different levels. One reason for this might be that articulating the relationship requires more than the resources dynamicism offers. In the dynamical approach, the main way of relating systems is by means of coupling. As Beer (1997) proposes, one can always focus on the coupled system, or decouple the systems and view one system as influencing the control parameters of another. This, however, does not make it clear how systems at different levels might be of different character.

Here the mechanical perspective can help the dynamical. To begin with, the analysis into parts and wholes gives initial sense to what is involved in going between levels. However, there is a second feature that is even more significant: wholes usually do different things than the parts that comprise them. A neural pathway may execute a cognitive function; individual neurons in the pathway emit spike trains, thereby carrying out one specific task needed for the function; and mitochondria in the neurons generate ATP needed for the neuron to produce the spike trains. One of the keys to moving upward among levels is to specify the manner in which the parts are organized so as to use the specific contribution of each part in a manner appropriate to the overall activity to be performed. The mechanical perspective thus can help elucidate the "somehow constructed" element in van Gelder's analysis.

However, there is at least an appearance of tension in promoting this union of the dynamical and the mechanistic. In contrasting the computational and the dynamical perspective, van Gelder associates concern with *state* and *structure* with the computational approach. He says, for example, "Computationalists focus on internal structure, and in particular on internal combinatorial or syntactic structure – how basic pieces are combined to form structured wholes" (sect. 4.2.3.2). In its concern for localization and organization, the mechanical approach seems to share the computationalist concern with structure and combining of pieces to form structured wholes. This appearance of unity between computationalists and mechanists, however, stems from an equivocation on the notion of structure. The mechanist is concerned with physical structure, but computationalists are concerned with formal structure – structure in symbol systems. The computationalist's long-standing lack of interest in the way computational systems are instantiated reveals a greater alienation between computationalists and mechanists.

The mechanistic and dynamical perspectives are hence natural allies, and the concern of the mechanist for such things as states and structures need not concern the dynamicist. A long-standing feature of the mechanist perspective is that one needs constantly to shift perspective between structure and function. When examining structure, one focuses on (temporary) stabilities; when focusing on function, one focuses on change. However, as soon as one decomposes the behavior of a structure, one is concerned with the activity within the structure, activity that can change the structure itself. Dynamics provides a set of tools for analyzing activity, but the identification of structures often provides guidance about the entities whose properties define the variables that change in value and whose patterns of change are to be analyzed in dynamical terms. (For an example of dynamical analysis that also invokes mechanistic decomposition in analyzing neural network robot controllers, see Cliff et al. 1996.)



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

## Framing the debate between computational and dynamical approaches to cognitive science

Randall D. Beer

*Department of Computer Engineering and Science and Department of Biology, Case Western Reserve University, Cleveland, OH 44106;*
**beer@alpha.ces.cwru.edu yuggoth.ces.cwru.edu/beer/beer.html**

**Abstract:** van Gelder argues that computational and dynamical systems are mathematically distinct kinds of systems. Although there are real experimental and theoretical differences between adopting a computational or dynamical perspective on cognition, and the dynamical approach has much to recommend it, the debate cannot be framed this rigorously. Instead, what is needed is careful study of concrete models to improve our intuitions.

Dynamical ideas are slowly but surely beginning to filter into cognitive science and there are signs that they may fundamentally transform the field. Because some of these ideas appear to challenge basic assumptions of a traditional computational perspective, it is hardly surprising that there is growing debate about the relative roles of computational and dynamical ideas in cognitive science. As with many ideological disputes, this debate is often marked by acrimony, misunderstanding, and outright dismissal. The goal of van Gelder's target article is to convince the skeptics that there is a substantive empirical claim behind those who seek to apply the concepts and mathematical tools of dynamical systems theory to the analysis of cognitive systems. Although I am extremely sympathetic to this overall goal and I applaud van Gelder for his valiant attempt to clarify the issues, I fear that his strategy gets him into more trouble than is strictly necessary to defend the dynamical approach.

Having the distinction between computational and dynamical systems turn on whether or not the state space and the time set of a system are "quantitative" simply does not work. The integers over which digital computers operate are nothing if not quantitative, and metrics can certainly be defined on integer spaces. In addition, dynamical systems can easily be defined over discrete state spaces (e.g., cellular automata), and computational descriptions of the symbol sequences generated by discretizing the output of continuous dynamical systems can be used to characterize their complex dynamical structure (Crutchfield 1994). Finally, what about analog computers, on which van Gelder's definitions are strangely silent? Sections 6.2 and 6.3 try to address some of these difficulties. For example, van Gelder insists that in a dynamical system distances must bear some "systematic relationship to system behavior" (sect. 6.2.2, para. 4). But how can you tell whether a system is bouncing around its state space because its dynamics are chaotic or because its behavior is based on nonmetrical "formal properties"? Although van Gelder would like to dismiss such concerns as mere nit-picking, it seems to me that they fatally undermine the very coherence of his nature hypothesis. As mathematical formalisms, both computation and dynamics are sufficiently broad that there is no empirical fact of the matter about which kind of system a cognitive agent *is*. Unfortunately, the debate cannot be framed with the mathematical precision to which the nature hypothesis aspires.

What the debate between computational and dynamical approaches to cognitive science is really about is which is the most insightful, explanatory, penetrating, and parsimonious stance to take toward a cognitive agent. There are very real conceptual, mathematical, and experimental consequences of adopting a dynamical versus a computational perspective on cognition. A computational approach is concerned with how an agent extracts, represents, stores, and manipulates information about its situation. In contrast, a dynamical approach is more concerned with the way in which the interaction between an agent's intrinsic dynamics and the dynamics of its body and environment unfolds into an observed trajectory of behavior with particular stability properties.

A dynamical perspective offers several potential advantages over a computational one. First, it explicitly incorporates the temporal dimension of cognition. Second, a dynamical perspective provides a broader theoretical playing field, because by making fewer a priori commitments, it demands an account of theoretical entities (e.g., representation) that a computational perspective simply takes for granted, and it encourages the consideration of alternatives to such entities. Third, it very naturally incorporates the growing realization that the behavior of an embodied and situated agent must be seen as arising from the ongoing interaction between its nervous system, its body, and its environment (Clark 1997). [See also Clark & Thornton: "Trading Spaces" *BBS* 20(1) 1997.] Fourth, it provides a natural language for reconnecting cognition with the brain processes that support it, with the noncognitive behavior that humans also exhibit, and with the "mere" adaptive behavior of simpler animals. Thus, it may provide a better framework for understanding the emergence of cognition in development and evolution. For all of these reasons, a dynamical framework holds the promise of providing a unified theoretical framework for the cognitive sciences. In contrast, the traditional computational approach to cognition can be almost fully characterized by the long list of things that are "somebody else's problem."

The claim that a dynamical perspective is the better one to take is essentially van Gelder's knowledge hypothesis. My only problem with the knowledge hypothesis is that it is not a genuine scientific hypothesis, at least not in the traditional sense of making an empirically falsifiable claim. At issue here are not experimentally testable predictions, but rather competing intuitions about the sort of theoretical framework that will ultimately be successful in explaining cognition. Simply put, computational and dynamical approaches make different pretheoretical bets about what features of cognition are more fundamental, about what sorts of mathematical models and tools will be most applicable to these features, and about how they will eventually fit together into an overall explanatory framework.

I certainly agree with van Gelder that this debate will only be resolved (or dissolved) by empirical investigation. Any battle between computational and dynamical ideologies must be fought on a case by case basis, grappling with the real experimental data on particular cognitive behavior. Indeed, as van Gelder has reviewed at length elsewhere (Port & van Gelder 1995), this battle is already being joined in many areas of cognitive science. In the meantime, I think that the careful study of concrete examples is more likely to clarify the key issues than abstract debate over formal definitions. In particular, the design and analysis of idealized model agents holds great promise in this regard (Beer 1997), and such work is beginning to engage cognitively interesting issues (Beer 1996). Such models can serve not only as intuition pumps (Dennett 1980), but also as the experiment pumps and mathematics pumps that advance us toward a unified theory of the mechanisms of adaptive behavior. There is nothing like studying birds or trying to build an airplane to cut to the heart of the debate about what can and cannot fly.

## Why the dynamical hypothesis cannot qualify as a law of qualitative structure

Nick Braisby, Richard Cooper,[a] and Bradley Franks[b]

*Department of Psychology, London Guildhall University, London E1 7NT;* [a]*Department of Psychology, Birkbeck College, London WC1E 7HX;* [b]*Department of Psychology, London School of Economics, London WC2A 2AE, England;* **braisby@lgu.ac.uk r.cooper@psyc.bbk.ac.uk b.frank@lse.ac.uk**

**Abstract:** Van Gelder presents the dynamical hypothesis as a novel law of qualitative structure to compete with Newell and Simon's (1976) physical symbol systems hypothesis. Unlike Newell and Simon's hypothesis, the dy-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

namical hypothesis fails to provide necessary and sufficient conditions for cognition. Furthermore, imprecision in the statement of the dynamical hypothesis renders it unfalsifiable.

Van Gelder opens up the prospect of a new paradigm for understanding cognitive processes with arguments that provide an interesting and important challenge to traditional characterisations of cognition in terms of symbolic computation. He argues that the dynamical systems hypothesis (DH) represents a novel law of qualitative structure (LQS), one rivalling Newell and Simon's (1976) physical symbol systems hypothesis (PSSH). Notwithstanding the utility and potential appropriateness of dynamical systems theory within certain areas of cognitive theorising, we find van Gelder's characterisation of the dynamical systems hypothesis as a law of qualitative structure unconvincing.

Van Gelder presents DH in opposition to what he refers to as the computational hypothesis: a hypothesis that is underpinned by the PSSH. However, perhaps the greatest strength of PSSH is its explicitness and hence its potential falsifiability. As an LQS, it provides the "*necessary and sufficient* means for general intelligent action" (Newell & Simon 1976, p. 116, emphasis added): consequently, it may be falsified by showing either that general intelligent action does not require a physical symbol system, or that it requires something in addition to a physical symbol system. As a candidate LQS for cognitive agents, it systematically describes cognition and differentiates it from noncognitive domains. To substantiate the claim that DH is an LQS, van Gelder must therefore specify a set of such necessary and sufficient conditions for cognition. His discussion provides no such conditions.

Trivially, DH in its bare form does not provide a sufficient condition on intelligent action, and van Gelder concedes this. He indicates that there are many dynamical systems (e.g., ones instantiated by the solar system) that fail to be cognitive. The critical feature of DH then, if it is to count as an LQS, is van Gelder's qualification that cognitive agents instantiate dynamical systems that are "causally organized at the highest level relevant to an explanation of cognitive performances, whatever that may be" (sect. 4.1, para. 4). Thus, the claim that DH is an LQS rests on demarcating this highest relevant causal level. Note that such a level is specified by PSSH in the claim that symbolic computation is necessary and sufficient for intelligent action, because this is the level at which processing displays properties of designation and interpretation. However, rather than specifying this level, van Gelder "simply takes an intuitive grasp of the issue for granted" (sect. 4); but failure to specify this level renders DH untestable. This would be less problematic were the highest relevant level of causal organisation guaranteed to be dynamical, but the claim that DH is an LQS is undermined by van Gelder's acknowledgement that not all systems are dynamical (sect. 6.1).

Specifying the highest relevant level of causal organisation could perhaps be aided by substantive constraints on what might count as a dynamical system (and hence the nature and number of systems that a cognitive agent instantiates). However, van Gelder concedes the "semiarbitrariness of systems" (sect. 3.1, para. 5): the boundaries between systems are unclear, and van Gelder provides no criteria for individuating systems.

A second possible approach to the specification of the highest relevant level of causal organisation might emerge from restricting the variables that can comprise a *cognitive* dynamical system. However, this route is compromised by van Gelder's failure to deliver a precise, formal definition of dynamical systems (e.g., defining them as systems whose state is defined in terms of continuous, and hence quantitative, variables, where the behaviour of those variables with respect to time is defined in terms of differential equations would be sufficiently precise to provide testable accounts of cognition). However, even if we accept van Gelder's account of dynamical systems, several questions arise concerning the nature of the variables that comprise those systems. Presumably not just any variable will be appropriate to the highest relevant level of causal structure. The issue of specifying such a level therefore reemerges as a question concerning the kinds of entity variables that may be "bound." Van Gelder concedes that applying dynamical notions to cognition will require generating variations on what model variables correspond to (sect. 4.2.1), but he offers no account of how this might be either approached or constrained. Presumably the question can only be addressed by considering the semantics of variables; in other words, "variable grounding" arises as a problem (by analogy with the problem of symbol grounding in classical artificial intelligence/PSSH: Harnad 1990).

The question of variable grounding raises the issue of precisely where the burden of explanation falls under DH. Although van Gelder claims that dynamical explanations depend critically on the mere fact of variables being quantitative, this disregards the central role that their qualitative or semantic differences – their grounding – plays in explanations. Unless van Gelder can differentiate precisely which aspects of dynamical explanations depend solely on quantitative factors, and which depend on semantic or qualitative factors, the status of DH as a knowledge hypothesis – as a set of tools for understanding cognition – remains unclear.

Our claims concerning the lack of criteria for individuating systems, variables, and the highest relevant level of causal structure might be taken to be addressed by van Gelder's response to the "not as cognitive" objection (sect. 6.7). He replies that in DH, cognitive agents instantiate "dynamical systems of *quite special kinds*" (sect. 6.7, para. 2; emphasis added). Unfortunately, van Gelder provides no non-question-begging clarification of exactly what might constitute such a "special kind." The claim that DH is an LQS – and hence, the nature hypothesis, that cognitive agents instantiate dynamical systems of a special kind – reduces to a supposed relation between two intuitively specified structures, a relation that is too imprecise to define a claim that can be empirically assessed.

In sum, both the nature hypothesis and the knowledge hypothesis are compromised by van Gelder's failure to specify the necessary constraints on an LQS. The resulting understanding of cognition that is provided by dynamical systems remains merely intuitive.

## The dynamical model is a Perceptron

Bruce Bridgeman

*Department of Psychology, University of California, Santa Cruz, CA 95064;*
**bruceb@cats.uscs.edu zzyx.ucsc.edu/Psych/psych.html**

**Abstract:** Van Gelder's example of a dynamical model is a Perceptron. The similarity of dynamical models and Perceptrons in turn exemplifies the close relationship between dynamical and algorithmic models. Both are models, not literal descriptions of brains. The brain states of standard modeling are better conceived as processes in the dynamical sense, but algorithmic models remain useful.

Are dynamical and computational models fundamentally different, or are they just different ways of describing systems that solve problems? Van Gelder's "dynamical systems are computable" objection (sect. 6.4) claims to deal with this question, but it treats dynamical systems as if they were real objects, rather than mathematical models of real-world machinery. In fact, systems that can be considered dynamic are routinely modeled as garden-variety algorithmic computational systems.

An example is van Gelder's Figure 1, which he considers to be an ideal example of a dynamical system. However, the figure looks strangely familiar. It turns out to be a two-layer Perceptron, slightly reduced in connectivity at the second layer, with a couple of recursive calculations added at the end. The model would have been instantly recognizable to Frank Rosenblatt, the developer of Perceptrons (Rosenblatt 1960). It is the mathematical model from



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

which parallel distributed processing (PDP) models are descended, the classical method for performing computations in a self-organizing system.

Rosenblatt was simplifying the known neurophysiology of the time to build models that could be analyzed without computers as we know them. He knew that real brain computations are analog (coincidentally, one of the hallmarks of dynamical systems according to van Gelder, sect. 3.3). In neurons, information is processed in graded postsynaptic potentials spreading across the surface of the cell's membrane. The membrane literally constitutes the space of van Gelder's definition. Each graded potential spreads from a synapse, decrementing exponentially with distance from the synapse. At the same time the graded signal degrades exponentially in time, so that only graded potentials that occur close together in time can interact. Thousands of such potentials might be interacting on the soma and dendrites of a single neuron at any one time. Everything matters for conduction of the graded signals: the ionic concentrations inside and outside the neuron, the diameter of the dendrite at each point, even the shape of the neuron. This is the only place in the brain where inputs from disparate sources are combined to make something new. If this is not a dynamical system, it is difficult to imagine what is.

But what happens next? The results of all this computation are read at the axon hillock: if the result is a depolarization to less than a threshold voltage, local channels activated by depolarization can open, and an action potential starts. The action potentials have nothing to do with information processing – they only carry signals from one place to another without changing them. Like the waves that carry FM radio signals, the amplitudes of these potentials are not important. It is the time between potentials, or the statistical structure of the spike train, or the phase relationships among potentials on neighboring axons that matters. All of these are also analog intervals of time, and they are prime targets for analysis as dynamical systems. Also, they are not essential to nervous system function – the intensively studied nematode *Caenorhabditis Elegans* has 302 neurons but no action potentials at all. The animal is small enough that everything can be done in the graded mode.

How are neural networks generally analyzed by brain theorists? Usually they are seen as algorithmic processing machines, and indeed such interpretations are generally useful. The dynamic system of the neuron is simplified to a digital summing junction (McCulloch & Pitts 1943) and simulated on a digital computer, because analog quantities can be approximated to any desired degree of precision in a long enough string of bits. Furthermore, Turing had already proved in the 1930s that an astonishingly simple machine, easily built from McCulloch-Pitts neurons, could compute any function that is computable. From there it is trivial to simulate a dynamic system in digital hardware, just as the dynamic theorists uses dynamic models to simulate a natural system. The choice of whether the best mathematical model consists of systems of differential equations (dynamic) or of Boolean algebra (algorithmic) is largely a pragmatic one.

Perhaps some of the confusion in simulating and understanding brain function comes from the concept of a state; this may be an area where the dynamical approach can make an important conceptual contribution. Philosophers are always discussing "states": of consciousness, of alertness, and so forth. However, seen from the brain that is doing the job of creating consciousness, the state is ephemeral. When we look inside the brain we see no states, only constantly fluctuating scintillations of graded potentials and quickly flashing action potentials. We record from electrodes and see everything in flux. If there is no change, there is no function.

Where do states arise in such an environment? Put simply, they do not – what seem to be states from the outside are processes on the inside. A dynamical analogy might be the vortex in water flowing from a kitchen sink. The vortex has observable properties such as direction of rotation, surface contour, and so on, but it exists only as a dynamic system. The hardware that supports it, the round hole in the sink, does not suggest the swirling flow of the vortex. The vortex exists only as a continuity of motion: it is a process, not a state. Looking at brain anatomy, we see only the hole in the sink, not the vortices that make the process interesting.

## Leaky virtual machines and the best of both worlds

Alan Bundy

*Department of Artificial Intelligence, University of Edinburgh, Edinburgh EH1 1HN Scotland, United Kingdom.* **a.bundy@edinburgh.ac.uk**
**www.dai.ed.ac.uk/staff/personal_pages/bundy/**

**Abstract:** The concept of virtual machine allows us to combine the dynamical and computational hypotheses in an investigation of cognition. Van Gelder explicitly rejects this approach, but not only does it allow us to use the modelling technique most appropriate to the task, it also opens up a new range of phenomena where these techniques interact.

***The importance of virtual machines.*** Van Gelder's target article fails to mention the concept of *virtual machine.* This is a serious omission because the concept of virtual machine accounts for much of the power of digital computers to model phenomena of all kinds, including cognitive phenomena. Using virtual machines, computer systems can be viewed at a series of levels, each level providing the machinery for implementing the next. Occasionally van Gelder gives us hints of this concept (e.g., in the discussion of embeddedness in sect. 5, or when he rejects the claim that dynamical systems will have to implement computational mechanisms in sect. 6.9).

However, van Gelder's ignorance of virtual machines is betrayed by his impoverished description of the "four fundamental requirements on a system to count as a digital computer" in sect. 3.2, para. 3. These requirements demand specific representational commitments, for example, time as discrete order, variables as discretely valued, which are certainly true of some of the virtual machines of which computer systems are composed; but these requirements are not necessary and are often violated, for example, when a digital computer is used to model a dynamical system.

***The role of dynamical systems in cognition.*** There is a growing consensus about the important role of dynamical systems as the virtual machines at the lower levels of cognitive agents. This arises from the successes of connectionism in modelling low-level cognitive phenomena, such as early perceptual and effector processes. It is also supported by the clear (although still crude) analogy between neural nets and the structure of the brain. However, connectionism has a much poorer record in modelling higher level cognitive phenomena, such as reasoning and decision making. Symbolic techniques have a much better track record here. Rule-based programs and explicit knowledge representation have been used to great effect for planning, diagnosing faults, proving theorems, understanding language, and so on.

The consensus in artificial intelligence is that dynamical systems will be used to construct the virtual machines on which symbolic processing will be based. Therefore, both the dynamical hypothesis and the computational hypothesis will have roles to play in our eventual understanding and modelling of cognitive agents. Therefore, we could partially accept van Gelder's knowledge hypothesis and restate it as *some parts of cognitive science can and should take dynamical form,* while claiming that other parts of cognitive science cannot and should not. Van Gelder specifically *rejects* this consensus. He claims at various places, (e.g., sect. 6.8) that dynamical systems can be used at *all* levels, and hence that symbolic techniques can be dispensed with altogether. Indeed, he identifies this as "One significant difference between the dynamical approach and PDP-style connectionism" (sect. 6.8, para. 2).

***Can van Gelder's hypothesis account for van Gelder's cognition?*** It is not possible to reject completely the possibility of adequate dynamical accounts of higher level cognitive processes. One


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

can only say that with our current experience of the modelling power of dynamical versus symbolic techniques, this seems *very* unlikely. To see just how unlikely, consider the cognitive processes that van Gelder must have gone through to write his article.

1. Formulating and refining his definitions of the dynamical hypothesis, the computational hypothesis, the digital computer, and so forth.
2. Marshalling his arguments in favour of the dynamical hypothesis and against the various objections to it.
3. Collecting examples of various successes and failures of different modelling techniques.
4. Deciding how to organise all this material into an article of the appropriate length, style, and so forth, while conveying the essential argument successfully.
5. Deciding which words to use to express the meanings he wanted to convey succinctly and simply, but without oversimplification.

This list of course, only scratches the surface of the processes involved. Now imagine doing any of this with dynamical systems. What quantitative variables should we use and what metrics define on them? What are the differential equations? Just to ask these questions exhibits the gulf between dynamical modelling tools and the thing to be modelled here. This particular task cries out for an intermediate virtual machine that would provide the symbolic representations and rules with which this kind of modelling is more easily conducted. Van Gelder would deny us this virtual machine.

***Leaky virtual machines.*** Unfortunately, van Gelder's uncompromising stance discourages investigation of a potentially fascinating aspect of a dynamical virtual machine for symbolic processing – it may be *leaky.* In computer science, we strive for a clean separation between a virtual machine and its underlying substrate. We want to think solely in terms of the virtual machine and not have to worry about its implementation. However, the initial experiments in building a dynamical virtual machine for symbolic processing have failed to achieve such a clean separation; the implementation keeps leaking through.

Consider, for instance, attempts to build logic-based reasoning systems in which the logical formulae are stored with a neural net. We may assert some properties of some objects by training the net on them. However, when we try to retrieve these assertions, the details will be mixed up. An object may be retrieved that corresponds to none of the objects stored, but whose properties are those that are most popular among them, that is, a kind of typical object. This emergent effect has profound implications for understanding human cognition – both its power and its potential for error. It can only be investigated by building hybrid systems that combine both dynamical and symbolic modelling techniques.

ACKNOWLEDGMENT

I thank Aaron Sloman for comments on this paper.

## What *is* the dynamical hypothesis?

Nick Chater and Ulrike Hahn

*Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom;* **n.chater@warwick.ac.uk**

**Abstract:** Van Gelder's specification of the dynamical hypothesis does not improve on previous notions. All three key attributes of dynamical systems apply to Turing machines and are hence too general. However, when a more restricted definition of a dynamical system is adopted, it becomes clear that the dynamical hypothesis is too underspecified to constitute an interesting cognitive claim.

Van Gelder claims that the dynamical hypothesis entails three key properties, but all three properties apply to Turing machines, the paradigmatic nondynamical system.

***1. Quantitative in state.*** "A system is quantitative in state when there is a metric over the state set such that behavior is systematically related to distances as measured by that metric" (sect. 3.3, para. 3).

This is true of a Turing machine. Define the following metric: the distance between two states is the minimal number of steps between them. The behavior of the Turing machine systematically relates to this metric (at each step, the machine will step to a neighboring state in this metric). This does not, of course, imply that all neighboring states are equally *accessible,* but this holds true for dynamical systems as well, where one cannot, for instance, simply reverse the direction of time.

***2. Quantitative state/time interdependence.*** "A system is quantitative in *time* when time is a quantity; that is, there is a metric over the time set such that system behavior is systemically related to distances as measured by that metric . . . *amounts* of change in state are systematically related to *amounts* of elapsed time" (sect. 3.3, para. 5).

This is also true of a Turing machine. The standard metric over discrete times (such that the distance between $t = m$ and $t = n$ is $|n - m|$). Plus the distance metric over space just mentioned will suffice. System behavior is again systematically related to time in this sense. Also, this metric is neither trivial, nor only occasionally or accidentally related to system behavior. Contrary to van Gelder's claims, the notion of computation embodied by Turing machines has central interest in the time course of computation: computational complexity theory (Garey & Johnson 1979) is a fundamental topic in computer science. Algorithms are evaluated not only in terms of *effectiveness,* but also in terms of *efficiency;* that is, questions are standardly evaluated not only in terms of computability but also in terms of tractability. This concern naturally carries through to computational accounts of cognition (e.g., Falkenhainer & Forbus 1989). Furthermore, within the framework of the computational hypothesis, there are models that have sought specifically to capture the time course of human behavior. Recent examples of this are Anderson and Matessa's (1997) production-rule system of serial memory, which seeks to model latencies or the careful evaluations of competing models of analogy with respect to response time predictions by Keane et al. (1994).

***3. Rate dependence.*** "Rates of change depend on current rates of change" (sect. 3.3, para. 6). As stated, this is a tautology, because it is not clear what separates "rates of change" from "current rates of change."

Van Gelder elaborates: "In these systems, variables include both basic variables and the rates of change of those variables" (sect. 3.3, para. 6). This seems completely mysterious, because we are given no analysis of what it is for a system to *include* a variable.

Van Gelder does note that "a *variable* is simply some entity that can change. . . . The *state* of the system is simply the state or value of all its variables at a time" (sect. 3.1, para. 1). From this it seems that state is just defined extensionally in terms of an arbitrary set of variables. If so, given any concrete object, we can define a system by a set of variables associated with that object and then define a new system including these variables and their rates of change. The latter system will be dynamical, according to the criterion of rate dependence. For any concrete object whatever (including the brain), at any level of analysis whatever, it seems that we can trivially satisfy the third criterion just by adding additional variables by fiat. So we seem to be no further forward.

What alternative analysis might be more appropriate? Van Gelder's Table 1 gives seven previous definitions of dynamical systems. Of these, 1 and 2 are tied directly to their physical realization, and hence not relevant in this more general context, whereas 5, 6, and 7 are trivially satisfied by Turing machines (essentially because Turing machines evolve deterministically over time).

However, consider definition 3 that a dynamical system is "a smooth manifold together with a vector field" (Casti 1993). Because this definition requires that the state space be smooth, the Turing machine *is* ruled out, because it has a discrete state space.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

In brief, definition 4 states that dynamical systems are *continuous* deterministic systems, but once we realize that this is the fundamental claim, then it is clear that the dynamical hypothesis is simply too underspecified to be of any interest.

The computational hypothesis does not *merely* say that the mind is discrete at a high level of analysis. Instead, it applies a theory of symbolic computation of enormous theoretical richness and practical power. However, the dynamical hypothesis *does* merely state that the system is continuous – it says nothing about how it works, aside from the trivial truth that it should be studied using the diverse tools of dynamical systems theory. In short, the dynamical hypothesis has the same status that a putative "discrete hypothesis" concerning the mind would have had before Turing, von Neumann, and development of digital, symbolic computation: that is, it would be almost completely devoid of substance.

## What might dynamical intentionality be, if not computation?

Ronald L. Chrisley

*School of Cognitive Computing Sciences, University of Sussex, Falmer BN1 9QH, United Kingdom;* **ronc@cogs.susx.ac.uk**
**www.cogs.susx.ac.uk/users/ronc**

**Abstract:** (1) Van Gelder's concession that the dynamical hypothesis is not in opposition to computation in general does not agree well with his anti-computational stance. (2) There are problems with the claim that dynamic systems allow for nonrepresentational aspects of computation in a way in which digital computation cannot. (3) There are two senses of the "cognition is computation" claim and van Gelder argues against only one of them. (4) Dynamical systems as characterized in the target article share problems of universal realizability with formal notions of computation, but differ in that there is no solution available for them. (5) The dynamical hypothesis cannot tell us what cognition is, because instantiating a particular dynamical system is neither necessary nor sufficient for being a cognitive agent.

Given van Gelder's concession (in sects. 6.3, 6.5, and 6.10) that he is not opposing computation in general, just digital computation in particular, I do not disagree with his main point. It is indeed an open empirical issue which kind of computation best characterizes natural cognitive agents, but he goes about stating this in a misleading way. Yes, "research into the power of dynamical systems is an interesting new branch of computation theory" (sect. 6.3, para. 2). However, with that considerable concession in mind, van Gelder should not have thought he was rejecting effectiveness; he was only pointing out that processes that are quantitative (at the "highest level") can be effective – effectiveness need not imply digitality. Nor should he have named the view he is opposing "the computational hypothesis" when it is really a specific form of digital computation that is his target.

Although van Gelder wisely avoids the antirepresentationalism that has been the focus of some recent dynamical criticisms of computational accounts of cognition, he fails to resist mentioning antirepresentationalism altogether (section 4.2.3.9). It is not only quantitative systems that can accommodate nonrepresentational aspects of cognition. For example, Brooks (1992) has famously rejected representations in the construction of mobile robots that behave intelligently in real time in the real world, yet his subsumption architectures are not quantitative; they are of the same kind as digital computational architectures. Perhaps it is right to reserve the term "computation" for processes that involve representations; but then there is a natural superclass of digital computation (let us call it the class of "digital machines") that stands in the same relation to digital computation as dynamical systems stand to dynamical computation. Despite the claims of those van Gelder cites in this section, there is no reason to believe that dynamical systems have any "nonrepresentational" advantage over digital machines.

A distinction should be made between two senses of the claim that "cognition is computation." According to one sense (call it the "opaque reading"), computation is whatever is described by our current computational theory, and cognition is best understood in terms of that theory. The "transparent" reading, by contrast, has its primary allegiance to the phenomenon of computation, rather than to any particular theory of it. It is the claim that the best account of cognition will be given by *whatever theory turns out to be the best account of the phenomenon of computation.* The opaque reading is a claim about specific theories, whereas the transparent claim is a claim about the phenomena of computation and cognition themselves. The "cognition is computation" claim can be true on the transparent reading, even if cognition is not best understood in terms of, for instance, formal operations, just as long as such operations turn out not to be good accounts of what makes actual computers work. I am one of those who believe formal notions of computation to be inadequate theoretical accounts of actual computational practice and artifacts (what Brian Smith [1996] has called "computation in the wild"). Van Gelder, however, insists (sect. 6.5) on opposing the formal notion of computation. This is understandable, because the formal view of computation is the de facto orthodoxy, and we are still waiting for a nonformal theoretical alternative. However, if it turns out that what makes the artifacts of Silicon Valley tick is not best explained in terms of formal computation, then van Gelder's discussion will have nothing to say against the transparent version of the "cognition is computation" claim.

Van Gelder's focus on formality in characterizing his opponent seems to have the unfortunate consequence of causing him to characterize dynamical systems as likewise formal. A recurring criticism of the computational approach is that its formality renders it universally realizable – Putnam (1988) and Searle (1990) argue that any physical system can be interpreted as realizing any formal automaton. This has the consequence that an account of cognition cannot be in terms of formal computation, because any particular formal structure whose realization one claims is sufficient for cognition can be realized by any physical system, including those that are obviously noncognitive. Dynamical systems as van Gelder characterizes them also seem to be universally realizable in this sense – one can use Putnam's tricks to show that every physical system instantiates every dynamical system. However, the difference is that there is a known way out of this problem for digital computation, whereas there is none for dynamical systems. Because computation is not purely formal but includes an implicit notion of discrete states and causal transitions between them, one can use this to restrict the set of physical systems that can be properly said to instantiate any given computation, thus avoiding universal realizability (Chrisley 1994). How are we to so restrict the set of physical systems that realize any given dynamical system, without rendering the dynamical system nonquantitative in the process?

Van Gelder's response to the "not as cognitive" objection (sect. 6.7) will not help him here. What he says is correct: just as the digital computation hypothesis does not claim that all digital computers are cognizers, but rather that cognizers are a special kind of digital computer, so also, mutatis mutandis, for the dynamical hypothesis (DH). The DH is not giving sufficient conditions for cognition. However, it does claim that the sufficient conditions can be given in terms of dynamical systems, as he has construed them, and the universal realizability points just made cast doubt on that.

Perhaps the universal realizability point can be countered for dynamical systems, as it was for digital computational systems. Nevertheless, there is a difficulty that arises out of van Gelder's admission that the DH is not providing sufficient conditions for cognition: it puts all the weight on the other foot. It implies that the theoretical value of the DH must be in its providing *necessary* conditions for cognition. However, van Gelder admits that the DH is not giving necessary conditions for cognition, either. Because the DH takes no stand on the nature of artificial cognition (sect. 4, para. 2), it is not a constitutive claim about the essence of cog-


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

nition in general, but rather a contingent claim about natural cognizers. Aside from relying on a natural/artificial distinction that removes us and our artifacts from the natural world, rather than seeing us/them as continuous with it, the DH has the drawback of leaving us without a constitutive account of cognition. The most likely place to look for such an account is not in the particularities of natural cognizers, but in the commonalities among all the systems worthy of the title: natural cognizers, natural cognizers in other possible worlds, and (as yet hypothetical) artificial cognizers. What do (natural) quantitative intentional effective systems and (artificial) digital intentional effective systems have in common? Intentional effectiveness. Perhaps, then, that is the true nature of cognition.

## Dynamical embodiments of computation in cognitive processes

James P. Crutchfield

*Physics Department, University of California, Berkeley, CA 94720-7300, and Santa Fe Institute, Santa Fe, NM 87501;* **chaos@santafe.edu**
**www.santafe.edu/~jpc**

**Abstract:** Dynamics is not enough for cognition, nor it is a substitute for information-processing aspects of brain behavior. Moreover, dynamics and computation are not at odds, but are quite compatible. They can be synthesized so that any dynamical system can be analyzed in terms of its intrinsic computational components.

It is hard to argue with the hypothesis that time underlies cognition. From our current vantage point, it is now simply odd that "static" conceptions of cognitive processes have held so much sway over alternatives during the last two decades, even at cognition's lowest levels. Models of early visual processing that view the flow of information only from the environment inward ignore the strong and numerous neural pathways from visual and higher cortex that reach back to early stages. Predispositions to such feedforward architectures essentially ignore time – time that is intrinsic to the behavior of neural systems and that supports the storage, transmission, and manipulation of information. Many similar examples provide enough justification to make a case for time underlying cognition.

One can conclude that time is one of the substrates of cognition. However, what kind of substrate is it? One of the central ways in which a natural process embodies time is in an architecture with feedback pathways between its components. Feedback in turn enables a system to exhibit a vastly richer range of dynamical behaviors than systems without it. That the temporal aspects of a process are well modeled by a mathematical abstraction called a "dynamical system" has been appreciated since the turn of the century. At that time, the French mathematician Henri Poincaré, through pure insight and hard hand calculation – that is, without the benefit of fast simulation on computers – discovered the essential mechanisms that lead to complicated and rich behaviors in dynamical systems. In dynamical systems theory, time is made implicit and one uses a geometric view of how structures in the state space generate as well as constrain behavior and the emergence of spatiotemporal patterns. In this way, kinds of temporal behavior are translated into geometric objects of varying topologies. As has been noted many times before, dynamics is the geometry of behavior (Abraham & Shaw 1983; Smale 1980).

Therefore, one can argue that, to the extent that it allows us to understand nature's complex nonlinear processes, dynamics is involved, at least partly, in cognition. However, it is an entirely different question to ask, How do dynamical systems support information processing and computation? This question leads to my main point: although I cannot argue against time and dynamics in cognition, any hypothesis connecting them to cognition has to be, and at this time can be, more concrete.

Dynamics is not a substitute for information processing and computation in cognitive processes. Given this, how can we synthesize dynamics and computation? The approach of computational mechanics is to analyze dynamical systems in terms of how geometric state space structures support computation (Crutchfield 1994a; Crutchfield & Young 1989). How much memory of the past is stored in the system's current state? How is this information transmitted between the system's degrees of freedom? How is it manipulated to produce the system's future behavior? What is the causal structure of this information flow? Building on a notion of the "effective" causal states embedded in a dynamical system, these questions can be answered both quantitatively and architecturally.

One consequence of such a synthetic framework is that it renders entirely moot both the disputes between symbolic and dynamical approaches to cognition and the debates about discrete versus analog embodiments of computation. In short, there simply is no dichotomy between a dynamical view and a computational view, as long as one is willing to assess fairly where each field is and to extend the notions each brings to the problems of cognition.

Are dynamical systems theory and our notions of computation ready to form a foundation for cognitive science? There are such fundamental limitations at present to both that I do not think so. Dynamics, for example, is in need of serious modification and extension to be made directly relevant to cognitive processing. During its century-long history, dynamical systems theory has been (1) deterministic, (2) low dimensional, and (3) time asymptotic. Neural and cognitive systems, in stark contrast, are (1) stochastic, (2) distributed and so high dimensional, and (3) must react quickly, that is over transient, not asymptotic time scales. The theory must be extended in just these ways.

Equally important, we need to greatly soften the digital hegemony that conflates "computation" with "discrete computation." The history of the dominance of mathematical logic in computation theory is well appreciated, but nature is structured in many other and quite different ways, as just noted. To the extent that these alternatives are the substrates out of which cognitive systems are built, we need to investigate those kinds of information processing supported by natural systems.

In summary, I agree that time is central to cognition. It is embodied in architectures that include feedback pathways. Feedback opens up qualitatively new regimes of dynamical behaviors. Currently, the best way to analyze the mechanisms that produce these behaviors is dynamical systems theory, which gives a geometric view of state space structures. However, to say that dynamics underlies cognition is not enough.

It is very important to distinguish the dynamical hypothesis from questions about how any given dynamical system supports computation – that is, how it stores historical information, transmits it internally, and manipulates it to produce its future behavior and output responses. One approach to this that avoids artificial dichotomies, such as symbolic versus connectionist or discrete versus analog embodiments, is computational mechanics. Computational mechanics allows one to identify the dynamical mechanisms that support intrinsic computation and so to lay out a system's embedded information-processing architecture.

Yet another different and higher level question is how the behavior of a dynamical system that supports intrinsic computation takes on functionality and cognitive import in an environment. This is the notion of intrinsic emergence, in which a dynamical system generates behaviors and patterns that take on functionality within itself, in particular without reference to a posteriori analysis by an external observer (Crutchfield 1994b). To my mind, this is one of the key open issues currently precluding a mathematical theory of cognition.

ACKNOWLEDGMENT

This work was supported at UC Berkeley by ONR grant N00014-95-1-0524.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

## The dynamical hypothesis: The role of biological constraints on cognition

Keith Davids and Simon Bennett

*Motor Control Group, Manchester Metropolitan University, Alsager ST7 2HL, Cheshire, United Kingdom.* **k.davids@mmu.ac.uk www.mmu.ac.uk**

**Abstract:** For the dynamical hypothesis to be defended as a viable alternative to a computational perspective on natural cognition, the role of biological constraints needs to be considered. This task requires a detailed understanding of the structural organization and function of the dynamic nervous system, as well as a theoretical approach that grounds cognitive activity within the constraints of organism and ecological context.

Van Gelder targets the current zeitgeist in cognitive science: what it means to be an intentional agent. Within a well-argued case against a number of general objections, the traditional computational view of brain function is eschewed for the systemic perspective of the dynamical hypothesis (DH). However, although van Gelder states explicitly that the appropriate domain of DH is "natural cognitive agents," surprisingly little is made of the failure of the computational hypothesis (CH) to consider the biological constraints on nervous systems. To be fair, van Gelder touches briefly on the lack of biological context in computational accounts of human behavior and also highlights the need for an interdisciplinary approach in proposing and defending DH. In this commentary, we focus on two key arguments in favor of the DH that van Gelder fails to substantiate in his target article. First, it is increasingly clear that, owing to the high level of interconnectivity in biological nervous systems, "fuzziness" rather than "digitality" best characterizes neural networks. Second, what is missing from computational accounts is any reference to the major constraint imposed by the biological context of behavior (Edelman 1987).

It is apparent that biological neural nets, exemplified by brains, are not static, deterministic machines, but dynamic, continuously fluctuating organs (Globus 1992; Kelso 1995). The simplified, silicon-grounded networks in an electronic computer are well designed for acting as digital communication channels with definite inputs initiated at clear points in time, symbolic transformations based on internalized rules, and outcomes that can be samples and reanalyzed. For natural cognitive agents to function digitally, they need to depend on static hardware components, reliable algorithms, and sequentiality of input, transformation, and output processes. These characteristics allow the very essence of digital computation, and exclude any possibility of "fuzziness."

What is the biological evidence for this hypothesis? In reality, the comparison with the hardware and software of electronic computers is inadequate. Biological nervous systems develop in highly variable ways. They are complex, "open" systems whose microcomponents are continually modifying and adapting their structural organization in response to a range of constraints. Nervous systems are better likened to the surface of a landscape undergoing continuous perturbations from internal and external pressures. Patterns of connectivity are not static and fixed but emerge between the molecules of organic matter continually shaped by chemicoelectrical constraints. Symbolic information storage does not sit well in such a dynamic, biologically based description of the nervous systems of cognitive agents. Instead, stabilities and instabilities, critical fluctuations, and pattern formation in the interconnected neural landscape under constraint favors DH for cognitive functions such as memories, perception, and attention (Kelso 1995).

Even the hardware components of the neural landscape are variable and emergent, being dynamically modified in the form of continuous cell migration, death, adherence, and differentiation. The dynamic nature of the constraints on biological nervous systems makes it virtually impossible to isolate digital inputs and to trace the symbolic transformation of information within the system. Real components of biological networks do not function digitally but are constrained by contextual parameters such as variable levels of excitation, fatigue, and connectivity in a dynamic, internal chemical milieu. It is perhaps unfortunate that, at one level of description, the brain does appear to resemble a kind of organic computing device. The on–off nature of neuronal activity, coupled with the seemingly point-to-point architecture, imply a digital character. In reality, the electrochemical communication between neurons connected directly (point-to-point) and indirectly (via the vast network of integrated feedback loops), involving states of polarization or depolarization, resembles a form of "pseudodigitality" in biological nervous systems. Clearly, despite the pervasiveness of the computational view of brain function, neurobiology has failed to provide reliable evidence for an architecture suited to symbol manipulation and syntactic communication in biological nervous systems (Daugman 1990).

The second serious flaw in CH, supporting the attraction of DH, is its lack of "embeddedness" in body and environment. If gaining "insights into the nature of people" is the overarching goal of van Gelder's arguments, surely a most important charge to be made against the CH regards the degradation of the role of environment in constraining cognitive functions. From a computational perspective, much care has gone into defining internal structures and hypothetical mental processes purported to underlie human behavior. However, the relationship of these cognitive processes to the typical environments in which cognitive agents perceive information for action is conspicuously ignored. Like van Gelder, we contend that natural intending agents are nothing when taken away from natural contexts (Gibson 1979). Computational accounts are "claustrophobic" in highlighting symbol manipulation, hypothetical internal representations, and mental processes in explanations of how intending agents acquire knowledge about their environments (Heft 1989; Reed 1993). Intentions toward environmental objects are to be understood in a biological context. They are not static, discrete and private affairs. They are characterized by goal-directed, striving, persistent physical activity in relation to environmental surfaces, objects, and events. For Reed (1993), intentions in natural cognitive agents are based in the "real world" and are constrained by mind, body, and social and biological contexts.

In conclusion, DH and its emphasis on constraints recognizes the failure of computational accounts to provide a biologically relevant explanation of intentionality. Intentions require information detection in natural environments, not symbol interpretation and the formation of private, internal representations of the world. Intentions are only one of a number of constraints on the behavior of natural cognitive agents. They can be embodied through the assembly of patterns of neurons depending on the numerous, instantaneous internal and external pressures on the cognitive agent. Although the multicausality of constraints in biological systems is understood in DH, computational accounts prefer explanations of single causality. It is becoming increasing clear that the shortcomings of CH are resulting in a paradigm crisis. We agree with van Gelder that the DH is well placed to act as its successor.

## Revolution, no! Reform, si!

Daniel C. Dennett

*Center for Cognitive Studies, Tufts University, Medford, MA 02155;* **ddennett@tufts.edu www.tufts.edu/as/cogstud.htm**

**Abstract:** Van Gelder's hard line against representations is not supported or supportable, and his soft line in favor of dynamical systems thinking as a supplement to representational models of cognition is good advice, but not revolutionary, as he seems to think.

There seem to be two van Gelders: the "Hard Line van Gelder" says, in effect, There are no representations at all, anywhere in the brain, in any useful sense – down with representationalism! The "Soft Line van Gelder" says, in effect, We must replace misshapen



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

ideas of representation – for example, Newell's (1980) physical symbol system hypothesis and other GOFAI (Haugeland 1985) ideas – with more perspicuous, biologically sound versions. The trouble with the hard line is that both good, working cognitive systems (brains) and defective, hallucinating cognitive systems are dynamical systems. The eschewing of all forms of representation does not permit the theorist to distinguish veridical or reliable perceptual systems from defective or ill-designed ones, since characterizing the difference depends on the use of some semantic (i.e., representational, in the broadest sense) notion or notions. Similarly, both a trained acrobat at his best and a drunk lurching down the road bumping into things are dynamical systems. You cannot begin to explain why some dynamical systems have evolved and others, just as dynamical, have not, until you introduce informational (or semantic) considerations of some sort that let you distinguish between when an animal is getting it (roughly) right and when it is getting it wrong. Once van Gelder responds to objections like this by retreating to the soft line, a lot of what he says makes fine sense, but then it is no longer revolutionary; it is simply a worthwhile call for taking dynamics more seriously in the context of good new-fashioned cognitive science. Indeed we should, but we don't have to get on any radical bandwagon to do this; just as well, since the bandwagon isn't going anywhere we want to go. Perhaps van Gelder didn't want us to read anything so radical into his essay, but if not, I am baffled by the strategy he adopts.

## All information processing entails computation, or, If R. A. Fisher had been a cognitive scientist . . .

Eric Dietrich[a] and Arthur B. Markman[b]

[a]*Department of Philosophy, Binghamton University, Binghamton, NY 13902;* **dietrich@binghamton.edu** [b]*Department of Psychology, Columbia University, New York, NY 10027;* **markman@psych.columbia.edu**

**Abstract:** We argue that the dynamical and computational hypotheses are compatible and in fact need each other: they are about different aspects of cognition. However, only computationalism is about the information-processing aspect. We then argue that any form of information processing relying on matching and comparing, as cognition does, must use discrete representations and computations defined over them.

Ordinarily, competition is a good thing. It might seem obvious that a lively debate between computationalists and dynamicists would benefit cognitive science, and that even if computationalism eventually won, making it fight for the title of "*the* foundational hypothesis of cognitive science" would improve both it and the field in general. However, in this case, competition is a bad thing, and what seems obvious is, in fact, wrong. We cognitive scientists will waste valuable time and money if we debate which hypothesis is correct, because the truth is that they are both needed for a theory of the whole brain and mind. In this commentary, we discuss this point and present a new argument that cognition is best viewed as computation over mental representations.

The computational and dynamical hypotheses seem to be related in two distinct dimensions that we will simply call the *vertical* and the *horizontal.* In each dimension, we contend, there is a right way and a wrong way to view the relation between the computational and dynamical hypotheses.

In the vertical dimension, the two hypotheses can be viewed as (1) competing at the same level(s) of description (i.e., the truth of one hypothesis entails the falsity of the other), and (2) both being true, but at different levels. The second, we claim, is the proper vertical relationship between the two hypotheses. Specifically, the computational hypothesis *supervenes epistemologically* on the dynamical one. That is, the computational hypothesis is a foundational hypothesis about cognition; the dynamical hypothesis is a hypothesis about the substrate of cognition: the behavior of the stuff cognition is implemented in and arises out of. Epistemological supervenience depends on *ontological supervenience.* The computational hypothesis supervenes epistemologically on the dynamical one precisely because computational *facts* about the *mind* supervene ontologically on dynamical facts about the brain and its neurons. Van Gelder discusses this relation between the two hypotheses in his objection (sect. 6.8). However, his discussion and rejection of this point is completely unconvincing. In section 6.8, he merely asserts without argument that the dynamical hypothesis is not limited to any level, certainly not to a level supporting computation. Of course, he can *define* the dynamical hypothesis so that it competes directly with the Computational Hypothesis (which is in fact what he has done). This makes for nice fireworks, but in the absence of a convincing candidate for a dynamical explanation of some aspect of cognition, fireworks is all it is.

The horizontal dimension is best thought of as an array of processes that includes sensory and motor functions as well as higher level functions such as language use and reasoning. Again, there are two ways to view how the two hypotheses are related. Van Gelder would have us believe that the dynamical hypothesis is adequate for explaining everything from the sensory and motor functions to the higher level ones. However, we think that the best way to view matters is to say that the dynamical approach is best for *some* of the sensory and motor functions, especially those with tight feedback loops, whereas the computational approach is best for higher cognitive processes as well as some perceptual-motor processes. Again, insisting as van Gelder does, in replying to the objection in section 6.7, that the dynamical hypothesis is intended to explain cognition as well as fingerwagging, is not convincing in the absence of an actual dynamical explanation of cognition (and we do not consider Petitot's attractor syntax [1995] to be a genuine cognitive explanation).

In short, in the cognitive agent, the dynamical hypothesis is best viewed as (1) being about the lower levels on which cognitive processing supervenes (the vertical dimension) and (2) being about processes that are peripheral to cognitive processes. We call our view the *compatibilist* view: the computational and the dynamical hypotheses are compatible because they are about different processes. At this stage, it is an open question whether the dynamical hypothesis is really better suited to just one of the dimensions, horizontal or vertical. This is the sort of thing we ought to be debating.

The situation we are urging van Gelder and like-minded dynamicists to avoid has transpired before in science. After the rediscovery in 1900 of Gregor Mendel's work on genetics, Darwin's theory of evolution via natural selection should have immediately been accepted, because Mendel's theory solved a long-standing problem about inheritance. However, the opposite happened: instead of cooperating, the "Mendelians" and the "Darwinians" battled it out for whose theory would win in the contest to explain the emergence of new species. Finally, after nearly 3 decades of pointless wrangling, biologists R. A. Fisher, J. B. S. Haldane, and S. Wright (with the help of many others; we are simplifying) showed that Mendelian genetics and Darwinian natural selection actually needed each other, and that both were true, being explanations at different levels. We think something like this is going on here with the computational and dynamical hypotheses, and we do not want to spend the next 30 years in pointless debate.

The analogy with early twentieth-century biology runs deeper than just the point that computationalists and dynamicists should cooperate. The problem of inheritance that Mendel's theory solved also allows us to make in a new and stronger way the familiar point that cognition requires discrete representations and computations defined over them.

During Darwin's lifetime, the mechanisms of inheritance were not well understood. The going theory, embraced by Darwin, was that inheritance worked by a sort of *blending* of whatever was responsible for offspring having the color, form, and other properties that they do. In 1867, Fleeming Jenkin, a Scottish engineer,



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

raised an objection to Darwin's theory of natural selection by pointing out that such blending inheritance would have the result that in a very short time, a population would become completely homogeneous, and then natural selection would have nothing to work on (short of the entire set of living things). Think of mixing a lot of different colors of paint; you quickly wind up with a sort of darkish sludge. Of course, Jenkin's objection was a reductio on the blending theory also, but this theory remained mainly because there was no other theory to replace it with. This is ironic because the rebuttal to Jenkin's objection was published at about the same time as Jenkin's objection. This was Mendel's *particulate* theory of inheritance; it was completely ignored until 1900.

The theory of blending inheritance assumed that the stuff responsible for inheritance was infinitely divisible. Fisher and the other population geneticists denied this, and worked out the mathematical details of a particulate theory that assumed that inherited properties are due to the summed small effects of a finite, but large number of particles. It is the relative frequency of discrete particles that are either present or not (genes) that matters. Pink flowers do come from red and white parents, but the process is not one of blending. Biologists now consider the discrete (digital) nature of genes a *necessary condition* for evolution by natural selection. This turns out to be a rather deep point about information, in general: *any sort of information-matching or information-comparing process must operate over discrete packets of information.*

Note that this point of ours logically precedes the powerful point of Fodor and Pylyshyn (1988) that combinatorial processes require discretely packaged information. Their point was about processes unique to cognition. Our point is about processes that *any* information-using system must have. The only way for van Gelder to avoid our point is to insist that thinking does not use information in any sense – a very implausible claim.

Computationalists and dynamicists alike claim that thinking consists of "information processes" in some sense. All cognitive processes, on anyone's theory, involve matching and comparing. What we learn from Fisher and colleagues is that any such process requires the information to be packaged in discrete chunks. Finally, any process that operates over discrete chunks of information is likely to be a computation. Therefore, at the level where *cognitive* scientists want to study them, the processes in humans are computational.

## The dynamical hypothesis in social cognition

J. Richard Eiser

*Department of Psychology, University of Exeter, Exeter EX4 4QG, United Kingdom;* **j.r.eiser@ex.ac.uk**

**Abstract:** Research in attitudes and social cognition exemplifies van Gelder's distinction between the computational and dynamical approaches. The former emphasizes linear measurement and rational decision-making. The latter considers processes of associative memory and self-organization in attitude formation and social influence. The study of dynamical processes in social cognition has been facilitated by connectionist approaches to computation.

The distinction introduced by van Gelder between the computational hypothesis and the dynamical hypothesis echoes debates in a field that he does not specifically discuss: attitudes and social cognition.

The computational approach in social cognition can be traced to the early influence of Thurstone (1928), who developed techniques to measure attitudes as interval scores on a unidimensional continuum of approval to disapproval. For practical purposes, he ignored the fact that different individuals could indicate the same levels of approval based on different reasons or associated thoughts. Positive and negative thoughts about an attitude object were effectively treated as canceling each other out rather than as separate facets that could coexist and be simultaneously activated (Cacioppo et al. 1997). Variation between contexts was taken to imply measurement error rather than a theoretically interesting form of situational influence. Attitude change was conceptualized as linear movement along the continuum of measurement. Expectancy-value models of decision-making and attitude-behavior relations (Ajzen & Fishbein 1980) provide a more recent example of a "Hobbesian" or computational approach to attitude theory. A core feature of these models is a formula to calculate an individual's overall evaluation of an attitude object from the sum of probabilistic beliefs about the consequences of an action, multiplied by how each consequence is evaluated.

In contrast, there was Allport (1935), who defined attitude as "a mental and neural state of readiness, organized through experience, exerting a directive and dynamic influence upon the individual's response to all objects and situations with which it is related" (p. 810). This is an explicit endorsement of the "Humean" or "dynamical" approach. Its contemporary expression can be found in research that regards attitudes as patterns of evaluative associations, stored in memory, the accessibility of which will vary as a function of contextual cues (Fazio 1990). New memory associations are constantly being acquired as one acts out one's preferences and receives feedback from the environment. Hence, attitudes, even where relatively stable, are never absolutely fixed. Their stability can be conceptualized as that of attractors within a dynamically changing energy landscape (Eiser 1994).

Although social cognition and attitude research has had its own version of the Hobbesian computational approach, this seemed to come about rather independently of any deliberate attempt to regard digital computers as valid models of the human mind. Although examples of "cybernetic" theorizing can be identified (Carver & Scheier 1981), by and large social psychologists regarded artificial intelligence as an alien culture (and no doubt the feeling was mutual!). As a consequence, social psychologists who are now starting to explore the potential of computational techniques for their own field are less encumbered by a commitment to the particular kind of computationality that is the target of van Gelder's criticisms. On the contrary, a major appeal of these techniques is that they offer a new possibility for examining *dynamical* processes of social cognition and interaction (Read et al. 1997; Smith 1996; Vallacher et al. 1994). A concept such as "spreading activation" – crucial to social cognitive accounts of associative processes in attitudes and impression formation – can be given a far more precise definition by considering what it would mean in the context of particular neural network architectures.

The related concept of self-organization preserves another Humean theme, that of the mind as a "republic or commonwealth" in which "higher" mental properties can emerge through the patterns of interaction between its constituent parts. It invites one to apply similar theoretical principles to groups of individuals who can communicate with and mutually influence one another. Again, computational techniques have proved invaluable in exploring the implications of theories of social influence, not least when considering large groups whose behavior would be difficult or impossible to observe in laboratory experiments (Eiser et al. 1998; Nowak et al. 1990). As van Gelder argues, the computability of such processes does not commit us to viewing cognitive agents as any less dynamical.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

## Dynamical models and van Gelder's dynamicism: Two different things

Chris Eliasmith

*Department of Philosophy, Philosophy-Neuroscience-Psychology Program, Washington University in St. Louis, St. Louis, MO 63130;*
**chris@twinearth.wustl.edu www.artsci.wustl.edu/~celiasmi/**

**Abstract:** Van Gelder has presented a position that he ties closely to a broad class of models known as dynamical models. One can support many of his broader claims about the importance of this class (as has been argued by connectionists for quite some time), but there are a number of unique characteristics of his brand of dynamicism. These characteristics engender difficulties for his view.

Van Gelder's article is commendable in many respects, but one's enthusiasm for certain aspects of his position is tempered by serious concern. A number of his central claims are compelling: of course we should consider cognitive systems to be dynamical; of course time is central to a good understanding of such systems; and of course we should use the tools of dynamical systems theory to help construct and understand our models. Do these ideas present a new direction that cognitive science has yet to embrace? Of course not. Many connectionists have embraced them for well over 10 years. However, there is more to the vision that van Gelder presents. I will argue that that "more" is precisely the problem with van Gelder's dynamicism.

First, to be more explicit about why this article is compelling and, in many ways, right: Unlike in his previous writings, van Gelder has seen fit to explicitly include connectionism in the class of (the right kind of) dynamical systems (sect. 1; in contrast, see van Gelder & Port 1995). Moreover, he rightly stresses the importance of modeling real time and of using our best mathematical tools to understand real time systems (sect. 3.3, para. 5 and 6; sects. 4.2.2 and 4.2.3). Hence one can agree completely that cognitive systems are dynamical systems. (Given the two previous points, we can derive a strong argument in favor of dynamical connectionist models.) Last, there are careful improvements over some of the more problematic claims in van Gelder's prior discussions of dynamicism. For example, his assertion that "the DH can embrace the idea that cognitive processes are computational" (sect. 6.3, para. 3), seems far more supportable than his previous suggestion that cognition is "the behavior of some (*noncomputational*) dynamical system" (van Gelder 1995, p. 358, emphasis added). Again, this view allies van Gelder with connectionists who have been making this suggestion for years (Churchland & Sejnowski 1992, p. 66). Overall, this target article presents van Gelder's most persuasive version of dynamicism (van Gelder 1993; 1995; van Gelder & Port 1995).

However, the problems that remain are central and, in my view, debilitating. Three will be outlined; the first is a problem of emphasis, the second a problem of interpretation, and the third, and most important, a problem in principle.

First, to place an emphasis on the discrete/continuous difference between classical computational systems and dynamical systems is somewhat deceptive in the context of real world cognition. There is strong evidence that the brain processes information in a manner that can be effectively captured by discrete descriptions. Bialek and Reike (1992) have estimated that individual spikes in neurons carry approximately three bits of information (see also Reike et al. 1997). This means that the timing of spike occurrences can be represented by a finitely precise set of values that easily could be given a discrete representation. If we really *needed* continuous variables in our models, then that would imply that an infinite amount of information was transmitted with each spike. This is highly improbable given the uncertain and noisy environments of neurons and brains. Thus, van Gelder's emphasis on the importance of the continuous nature of dynamical systems is uninteresting, if not misleading (sects. 2, 3.2, and 3.3).

Second, van Gelder's position has serious difficulties regarding model interpretation. How can we tell what a variable means? Barton (1994) discusses the difficulties of using such models in psychology, noting that they are strictly metaphorical. Although the metaphor's source is mathematical, it does not supply the model itself with any technical rigor. When van Gelder states, "Dynamical theories of cognitive processes are deeply akin to dynamical accounts of other natural phenomena such as celestial motion" (sect. 6.6, para. 1), he is ignoring an important difference between the two: we can measure angular velocity, but we have no idea what "valence" *is* in the decision field theory model or how to measure it in the real system. This difficulty stems from van Gelder's insistence on high-level and low-dimensional dynamical models (sect. 4.1), an insistence strangely contrary to admitting connectionist models into the class of the "right kind" of dynamical systems (sect. 1).

Third, and most important, van Gelder's version of dynamicism seems to sidestep a central issue in cognitive modeling – representation (Eliasmith 1996). The computational hypothesis (CH) cannot be so easily dismissed, particularly when compared to the (DH) dynamical hypothesis. The CH offers built-in prowess in the manipulation of symbolic representations. Of course, as van Gelder is quick to point out, this comes at a cost. Classical models have difficulty with many of the sorts of behaviors that dynamical models handle very gracefully. However, people are both dynamical *and* competent symbol manipulators. It would not be in the best interest of cognitive science to simply trade one set of strengths for another. Rather, a set of principles that is able to capture both sorts of behavior should be sought. It seems, in other words, that van Gelder has neglected the really interesting, and difficult, question of how a dynamical system *could* give rise to representational behavior and thus human-style intelligence.

In sum, claiming that cognitive systems are dynamical seems fundamentally right. However, it is fundamentally wrong to emphasize continuity, proclaim low-dimensional models superior, and all but dismiss the importance of representation to understanding human cognition. Dynamical models are clearly desirable. Van Gelder's dynamicism is another matter.

ACKNOWLEDGMENT

This research was supported in part by the Social Sciences and Humanities Research Council of Canada.

## Interlevel connections and agent evolution should not be overlooked

Donald R. Franceschetti

*Department of Physics and Institute for Intelligent Systems, The University of Memphis, Memphis, TN 38152;* **dfrncsch@memphis.edu**

**Abstract:** A consideration of underlying neural dynamics and the evolutionary process producing cognitive agents should complement the development of dynamical models of behavior. The geometrical aspects of dynamical systems theory which make it useful in the description of systems acting in an environment are less useful in understanding agents interacting with a range of environments, and may call for supplementation by evolutionary insights.

The key element of van Gelder's formulation of the dynamical hypothesis, or at least the one that calls for extensive confirmation by observation, is his assertion that the dynamical systems are "instantiated by the agent at the highest relevant level of causal organization" (sect. 4.3). Thus, the fact that this behavior is implemented by a network of neurons and that most neuroscientists assume that behavior at this microstructural level implements a dynamical system does not ensure the validity of the hypothesis. Likewise, the fact that a sufficiently large digital computer (and digital computers, with their input and output data are dynamical systems, albeit rather far from prototypical ones) can simulate, to



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

any desired extent, the behavior of a dynamical system does not mean that an identifiable dynamical system at the cognitive level will itself look like a digital computer. The assertion is that for most aspects of behavior, an appropriate dynamical system can be found, and the various dynamical systems will generally not look like digital computers in action, nor will they necessarily resemble each other. For van Gelder, the evolution of agents that instantiate dynamical systems is not of immediate interest, and one of the attractive features of the hypothesis is that dynamical systems are described by geometric spaces so that possible behaviors can be characterized in geometric language.

Although there is no fundamental logical problem with the notion of different dynamical systems for different aspects of behavior, the history of physical theory suggests that there is much to be gained by understanding how behavior at the macro or behavioral level is generated by dynamics at the micro level. Classical mechanics, after all, provided a very accurate theory of the dynamics of rigid and nonrigid solid objects and of liquid and gaseous bodies long before there was any quantitative understanding of the atomic nature of matter. On the other hand, the classical mechanics of macroscopic bodies provided no explanation of the existence of materials with rigid or elastic or fluid properties, or for the phase transitions that converted solids into liquids and liquids into gases at precisely defined temperatures. Furthermore, experimentation at the macro level kept turning up mysterious phenomena, like superconductivity and superfluid behavior, which could not be understood at all until an accurate theory of behavior at the micro level was obtained.

A key idea in connecting behavior at the macro and micro levels was that of phase space, which provided a geometric representation of all the possible states of the system. As Poincaré had pointed out near the end of the nineteenth century, this geometrical feature made it possible to draw meaningful conclusions about the possible trajectories of a classical system, even if the equations of motion were intractable from a mathematical standpoint. An appealing feature of the phase space formulation was the notion of basins of attraction, localized regions of phase space, such that once a trajectory entered one of them it would never depart. It is this feature that can perhaps best accommodate Hume's notion (as stated by van Gelder in sect. 1) of a "universal force of association, whereby ideas attract in proportion to their similarity," if the families of trajectories representing agents with each of the associating ideas ended in basins representing states with all of the ideas realized.

In terms of interactions with the environment, dynamical systems can be of three types. In the typical case considered in physics, the environment appears only in the form of time-independent parameters. In the second type, which might well be called driven systems, the environment appears as a time-dependent parameter and the natural phase space of the problem includes a dimension associated with the value of this parameter (Baker & Gollub 1990). In the most general case, the system interacts with its environment, with the system's state effecting changes in the environment's state and vice versa. Although it is always possible to view the system variables as the coordinates of a point moving in a space restricted to system variables only, much of the ability to generalize about trajectories is lost because many different states of the overall system will coincide in the system-only space.

Van Gelder's first examples of a dynamical system in psychology, the decision field theory (Busenmeyer & Townsend 1993), is of the first type, in that the inputs do not change as the system evolves to a decision. The second example (Kelso 1955) is arguably of the second type, with the required "wagging speed" as the external parameter. However, important aspects of behavior involve continuous response to and effect on the environment and would seem to fall inevitably in the third category. If the changes involved are not too rapid, it appears reasonable to think of the trajectory of the agent as evolving within the system-only space, with transitions induced by the environmental input as suggested by Horgan and Tienson (1996). With stronger interactions this may not be a useful procedure. Although one can consider the agent and environment as a single dynamical system, this overlooks the fact that the agent has evolved or learned to interact with a variety of environments. This suggests that the dynamical view will need to be augmented by evolutionary insights and perhaps a generalized form of the dynamical theory developed for the general case.

## The dynamical hypothesis: One battle behind

Robert M. French[a] and Elizabeth Thomas[b]

*Psychology Department, University of Liège, Liège, Belgium;*
**rfrench@ulg.ac.be www.fapse.ulg.ac.be/Lab/Trav/rfrench.html**
[b]*Institut Léon Frédéricq, University of Liège, Liège, Belgium;*
**ethomas@ulg.ac.be www.fapse.ulg.ac.be/Lab/Trav/ethomas.html**

**Abstract:** What new implications does the dynamical hypothesis have for cognitive science? The short answer is: none. The target article is basically an attack on traditional symbolic artificial intelligence (AI) and differs very little from prior connectionist criticisms of it. For the past 10 years, the connectionist community has been well aware of the necessity of using (and understanding) dynamically evolving, recurrent network models of cognition.

Our views on the current target article by van Gelder are nicely summed up by paraphrasing the opening lines of another recent article by him (1998) about connectionism and the philosophy of mind: What new implications does the dynamical hypothesis have for cognitive science? The short answer is: None. In what follows, we will sketch out the longer version of this answer.

Van Gelder's target article is a clarion call for the old and deeply flawed computational hypothesis (CH) and the establishment of a new basis for modeling cognition – the dynamical hypothesis (DH). However, the battle van Gelder is waging was waged (and largely won) a decade ago. Connectionism *really did* usher in a new era, a radically different alternative to the prevailing paradigm based on the physical symbol system hypothesis. One can quibble about some of the contributions of the connectionist revolution, but the fact remains that the advent of connectionism demanded a qualitatively different way of modeling cognition. Reading van Gelder's article, one has the impression that he believes the DH is another such sweeping, radically different way of viewing cognitive modeling. Unfortunately, it is not.

To show this, we suggest a very simple experiment. Download a copy of this target article into your favorite word processor. Then replace every occurrence of "the dynamical hypothesis" with "connectionism," of the words "dynamical" and "dynamicist" with "connectionist," of "dynamical systems theory" with "connectionist theory," and so forth. Then print out a copy of the modified paper. You will have a paper that, basically, could have been written around 1990, once the necessity of recurrent networks for modeling cognition was clearly established within the connectionist community.

The point is that van Gelder's article reads almost exactly like earlier connectionist attacks on the old symbolic (AI) paradigm – what he has rebaptized the computational hypothesis – and adds very little of significance to those criticisms. The author himself points out that connectionism "models cognition as the behavior of dynamical systems," (sect. 1, para. 4) referring to a major article by one of the leading members of the connectionist research community (Smolensky 1988). Elman (1990) published an important article that emphasized the recurrent (dynamical) aspects of connectionist modeling. In short, by the late 1980s, the importance of dynamics in connectionist modeling was well understood. That researchers in symbolic AI are not particularly concerned with dynamical modeling is beside the point. We suggest that van Gelder would have had a far more difficult time convincing people of the novelty of his ideas had he contrasted them with the


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

decade-old principles underlying research on recurrent connectionist networks.

Van Gelder continually emphasizes the need to apply the tools of modern dynamical systems theory to models of cognition. This is certainly sound advice. According to the author, the use of the DH framework provides a viable empirical alternative to the CH. As he correctly points out, investigators in cognitive science currently only use the tip of the dynamical iceberg. What he does not say is that this may have to remain the case for a very long time. Indeed, it may never be possible to make use of much of dynamical systems theory in studying cognitive (or even biological) processes. These theories were originally developed as mathematical models and applied to highly controllable physical systems, for example, electrical circuits. Biological systems, on the other hand, contain large amounts of noise and, more important, they have a high degree of nonstationarity. Nonstationarity refers to the lack of constancy in the laws governing the evolution of the system. Many of the measures of dynamical systems on which much of dynamical systems theory is based require very large amounts of data (Rapp 1993). Cognitive systems, however, tend to move rapidly from one state to another, often making it impossible to collect the quantity of data required to appropriately apply dynamical systems theory. However, even allowing for these problems, the target article goes far beyond merely advocating new tools to analyze recurrent systems; it is a call for the establishment of an entirely new modeling paradigm in cognitive science.

The one area in which our word-replacement experiment fails is in the discussion of anti-representationalism. The debate concerning the ultimate necessity of representations is an important one. However, it is hard to understand what van Gelder means when he says, "Within the dynamical approach, such systems [devoid of representations] can be not only imagined, they can be modeled and constructed" (sect. 4.2.3.9). However, these dynamical systems rely on the emergence of attractors, and what are attractors other than internal representations with a new name? Is van Gelder attempting to make some principled distinction between patterns of activation that persist in time (internal representations à la D.O. Hebb) and attractors? Antirepresentationalists bear the responsibility for clarifying this distinction, if indeed it can be done. [See also Amit: "The Hebbian Paradigm Reintegrated" *BBS* 18(4) 1995.]

Particularly puzzling to us is what van Gelder calls "False Objection 3 – Dynamical systems are computable" (sect. 6.4). Does his reply to this objection translate into the following claim: Even though a digital computer could, at least theoretically, simulate to an arbitrary degree of precision any analog function necessary for cognition, nonetheless, no digital computing device could ever be a cognitive agent? Even though functions that are "effectively computable" might *theoretically* require something more powerful than a computer (read: Turing machine, in this case) to compute them, the burden of proof lies with van Gelder to tell us (1) what those functions might be and (2) why they are important to cognition. He has done neither.

## Why dynamical implementation matters

James W. Garson

*Department of Philosophy, University of Houston, Houston, TX 77204-3785;*
**garson@menudo.uh.edu**
**bentley.uh.edu/philosophy/garson/Jim—Garson.htm**

**Abstract:** Another objection to the dynamical hypothesis is explored. To resolve it completely, one must focus more directly on an area not emphasized in van Gelder's discussion: the contributions of dynamical systems theory to understanding how cognition is neutrally implemented.

Suppose a researcher in the dynamical paradigm models a cognitive process as a set of nonlinear dynamical equations (with no known solution). The only practical way of understanding what this model says about behavior is to simulate it with a digital computer. (The alternative, namely, to use an analog machine, poses problems of inaccuracy and poor reproducibility of results.) Therefore, the only information about the dynamical model that the researcher can ever hope to obtain comes from a function that is digitally computable. It follows that there is always a model of the cognitive process being explored that meets conditions 1–4 of section 3.2 and so counts as classical. (To see why, note that the real-valued cognitive variables of the dynamical model can be recast in limited precision, hence digital form [condition 1]. Furthermore, those variables can be given the interpretations for the corresponding variables in the dynamical model [condition 4]. Variables for time are discrete [condition 2], and of course there is an effective algorithm [condition 3].) It follows that at least the knowledge version of the dynamical hypothesis (sect. 4.2) is in doubt. How can a dynamical model be more revealing than a classical one when the only behavior the dynamical researcher could possibly encounter is already classical? Because the classical approach is better entrenched, why not carry on with classical models for cognition?

Van Gelder mentions a line of rebuttal to this objection in section 6.4. The fact that a system's behavior can *be computed* classically does not show that the system is *itself* a digital computer, any more than the classical computation of a planet's orbit shows that the planet is a digital computer. Fair enough. However, the reply still leaves an open question. We can explore planets to verify that they do not compute their orbits, but how do we evaluate the corresponding issue in the case of cognition? Presuming we have two models – a dynamic one, and the corresponding digital model producing for all we know exactly the same behavior – what purchase do we have on the claim that one is a superior account of cognition? What else can be brought in to determine which model captures the true causal nature of the cognitive process?

The answer takes us in directions orthogonal to a trend in van Gelder's discussion. To prepare the way for a rebuttal to the wrong level objection (sect. 6.8), he has stressed (in sect. 4.1) that dynamical models of cognition contain "cognitive-level" variables, not characterizations of neural properties. This may leave the reader with the impression that dynamical cognitive science offers us a host of separate cognitive-level models for various cognitive functions, leaving neuroscience in the dust. On the contrary, one of the most exciting prospects offered by dynamical systems theory is in providing powerful new concepts of how cognition is implemented in the brain. The brain clearly has a dynamical description at the neural net level. Results concerning self-organized emergent structure in such dynamical models promise to provide a whole new battery of novel information-processing tools that could have been exploited by evolution to build a mind. Van Gelder mentions some of the relevant work in sections 6.9 and 6.10. However, he does not make it clear how important the development of successful dynamical implementation theories will be to a defense of the dynamical paradigm.

In a cognitive science that lacks a worked out implementational theory, it is hard to explain why we should prefer dynamical models over classical ones. Faced with dynamical models and corresponding classical ones with equal success at simulating a given cognitive function, we need some further way to characterize whether or not symbolic states are implicated in the causal structure of the processing. If the dynamical hypothesis is to be given explanatory priority, a successful implementational theory based on nonclassical structures will be needed.

One reason we no longer accept Ptolemeic epicycles in astronomy despite the fact that epicyclic models match the predictive success of Newtonian ones is that we have no "implementational theory" for epicycles. Given what we know about atoms, molecules, and physical matter in general, there is just no room for gigantic rigid rings and pivots that could force the planets to trace out their orbits. We now know that there is no need for such a mechanism. Orbit following "comes naturally"; it is dynamical or-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

der that emerges in the planetary context from laws of motion and gravitation.

The classical view presumes that cognition will require mechanisms resembling those found in digital computers. Implementation means showing how a symbolic processor and its software are constructed in the brain's neural net. What gives the dynamical hypothesis its most exciting explanatory potential is that it provides a much richer palette of possible implementation strategies for cognition. For example, analyses of neural nets trained on simple language-processing tasks (Elman 1995; Pollack 1991) have shown that solutions to the problem can be achieved by tuning the model to develop the right class of system trajectories. These nets appear to co-opt self-organized structures that are natural features of certain dynamical systems for use in cognitive processing. Methods of "representation" and "processing" that emerge do not leave room for the classicist's concatenated symbols with causal role.

Because the fundamental dispute between classicists and dynamicists is whether the causal story that underlies cognition involves symbols, and because the presence or absence of symbols depends on how cognition is implemented, the differences between classical and dynamical hypotheses cannot be successfully formulated without considering the differences between classical and dynamical theories of implementation.

## Cognitive dynamics: A psychological perspective

Richard A. Heath

*Department of Psychology, University of Newcastle, Callagan, NSW 2308, Australia;* **heath@psychology.newcastle.edu.au**
**psychology.newcastle.edu.au/~heath/**

**Abstract:** Although cognitive psychology is still dominated by computational theories, there is an emerging emphasis on dynamical aspects of cognition. Examples are provided supporting the increased use of dynamically inspired models by psychologists. Despite measurement and model verification problems in the direct use of dynamical system theoretic technology, van Gelder's general approach to cognition is recommended.

The distinction between the computational (CH) and the dynamical hypotheses (DH) about cognition is an important one. Of particular interest is the contrast between a cognitive theory that is time independent versus one emphasizing the temporal evolution of information processing. For example, some models of cognitive development propose that humans progress through successive stages of increasing information-processing complexity. The cognitive capabilities characterizing each developmental stage depend on theoretical concepts, such as an increase in memory capacity with age (Halford et al. 1994). The puzzle that remains is how cognitive processes evolve in this apparently stage-dependent manner. Van Geert (1994) proposed a dynamic developmental model that satisfies several of van Gelder's DH criteria. These include cognitive state changes represented as trajectories in a geometric space. Consistent with section 4.2.3.3 of the target article, the dynamic representation of cognitive development is "seen as the simultaneous, mutually influencing unfolding of complex temporal structures." Of interest is van Geert's theory of "connected growers," in which stagelike cognitive processes emerge spontaneously from a coupled nonlinear dynamical system. It is difficult to see how such a process could be modeled using the CH.

Cognitive psychology has been influenced by computationalist concepts, for example, the account of semantic categorization in terms of static hierarchical representations (Collins & Quillian 1969). In this example, response time (RT) data are used to justify the hierarchical structure rather than to evaluate the dynamic processes involved in memory retrieval. Because most experimental work in cognitive psychology involves observing the effect of independent variables, such as word frequency and word length, on dependent variables, there is little opportunity for the quantitative effects of cognitive dynamics to be revealed. Behavioral indices such as response accuracy and RT are frequently considered as equivalent, with both short RTs for correct responses and high response accuracy being associated with an increase in cognitive efficiency. Many researchers are blind to the possibility that a more thorough analysis of RT data might provide valuable insight into cognitive dynamics. Notable exceptions include the work of Ratcliff and McKoon (1982) and Dosher (1984), who limited processing time to facilitate the study of cognitive dynamics. For example, when subjects are required to respond within half a second, they mistakenly identify a *butterfly* as a member of the category *bird.* When accorded more processing time, the subjects no longer commit this error. Humans concentrate on the fact that both *butterflies* and *birds* fly and have wings, initially. It takes longer to determine that a *butterfly* is an *insect.* Expressed dynamically, the cognitive trajectory initiated by the stimulus *butterfly* lies initially in the vicinity of the *bird* attractor, but as time progresses it approaches the *insect* attractor more closely in some abstract cognitive space. It is difficult to see how the CH representation of semantic memory could possibly provide such insight into cognitive processing mechanisms.

Perhaps the most enlightening feature of the dynamic approach is using dynamical systems theory (DST) to answer the challenging questions in cognition such as: How does cognitive competence evolve over time within an individual and within a species? Why is it desirable to have a limited memory capacity? How do we learn a complex psychomotor skill, such as playing a piano? With the advent of sophisticated extensions of DST (Kauffman 1993; Langton 1990), we are beginning to understand the mechanisms that subsume such self-organizing, emergent phenomena including the pervasive nature of universal computation (sects. 2, 4, 5, 6.3).

A complication to van Gelder's suggestion that the DH permits quantification in cognitive theory (sect. 3.3) arises when one tries to quantify cognitive products that cannot easily be represented as continuous variables (as response accuracy and RT can). Rating scale judgments, for example, require ordinal measurement theory for which there is no simple dynamic correlate. In such applications, van Gelder's assertion that "every cognitive process unfolds in continuous time" (sect. 5, para. 4) may require extension.

A critical aspect of cognitive dynamics (sect. 6.1) is the influence of nonlinearity, which is difficult to intuit in a computational approach, especially when the dynamical system is close to chaos. In this situation, considering symbolic dynamics may provide a link between the computational and dynamic approaches to cognition. In so doing, a mapping from the DST geometry (sect. 4.2.3.2) into a formal linguistic structure might be feasible.

The distinction between serial (CH) and parallel (DH) processing (sect. 4.2.3.7) as useful cognitive theory is compromized by problems in identifying parallel, serial processes and their myriad combinations (Townsend & Ashby 1983). Finally, van Gelder's conclusion emphasizes the importance of "sustained empirical investigation" (sect. 7, para. 2), but there is little guidance on how such investigation can determine the relative validity of DH and CH. It may be the case that it is very difficult indeed to provide the empirical evidence needed to reject CH in most cognitive scenarios, using tools available to experimental psychology. Van Gelder's ideas should encourage cognitive psychologists to compare and contrast the two approaches to generate empirically sound cognitive theory, which may turn out to be domain specific. Nevertheless, the dynamical approach is immensely promising.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

## Resisting the tyranny of terminology: The general dynamical hypothesis in cognitive science

Terence Horgan and John Tienson

*Department of Philosophy, University of Memphis, Memphis, TN 38152;*
**thorgan@memphis.edu**
**www.people.memphis.edu/~philos/horgan.tienson**

**Abstract:** What van Gelder calls the dynamical hypothesis is only a special case of what we here dub the general dynamical hypothesis. His terminology makes it easy to overlook important alternative dynamical approaches in cognitive science. Connectionist models typically conform to the general dynamical hypothesis, but not to van Gelder's.

Cognitive science is usefully viewed as using three interrelated levels of description: (1) a "top" level, at which states and state-transitions of a cognitive agent are described in psychological vocabulary; (2) a "middle" level, at which they are described in a purely formal/mathematical way, apart from any informational content; and (3) a "bottom" level, at which they are described physically. The interlevel relations are what computer scientists call "implementation" and philosophers call "realization": cognitive states are implemented by mathematical states, which are in turn implemented by physical states. Classical cognitive science assumes that the middle level involves the discrete mathematics used in the theory of computation: rules for manipulating symbol-structures on the basis of their formal-syntactic properties. However, in an alternative, nonclassical, approach, the middle level is instead assumed to involve the (fundamentally continuous) mathematics of dynamical systems theory. We will call this the general dynamical hypothesis.

A special case of this hypothesis arises by collapsing together the top and middle levels: view cognition as describable by a relatively low-dimensional dynamical system $D$ such that (1) each dimension in $D$'s state-space corresponds to some psychological variable, and (2) $D$ is characterizable by means of a set of differential equations over these psychological variables. We will call this the top-level dynamical hypothesis.

Van Gelder's discussion, in the target article and in other of his writings, strongly suggests that what he calls the dynamical hypothesis in cognitive science is the top-level hypothesis rather than the general one. He is concerned in the first instance, he says, "with how agents are causally organized at the highest level relevant to an explanation of cognitive performances, whatever that may be." (sect. 4.1) However, beware of the tyranny of terminology! Appropriating the generic label "the dynamical hypothesis" for what is really just a species of a wider genus of potential dynamical approaches in cognitive science threatens to obscure other important possibilities.

Connectionist models often fall within the scope of the general dynamical hypothesis, but not within the scope of the top-level dynamical hypothesis. Explaining why a given connectionist model works will typically involve reference to all three levels of description. At the top level are psychological states that are generally not quantitative magnitudes (and whose behavior is hence not in general describable by differential equations over psychological variables). At the middle, mathematical, level is a dynamical system – that is, for an $n$-node network, an $n$-dimensional state-space determined by the activation values of the respective nodes – together with a collection of temporal trajectories through this state-space that jointly constitute an $(n + 1)$-dimensional "activation landscape" for the network ("downhill" being the direction of time). At the bottom level are physical states, each of which is a total-activation state of the network whose temporal-transition profile is described by the dynamical system.

Explaining how a given connectionist model successfully accomplishes the task it was trained to perform will typically appeal simultaneously to two interconnected factors: (1) topographical properties of the activation landscape (e.g., the number of attractor basins, the structure and positioning of basin boundaries); and (2) facts about the way psychological states are implemented mathematically as points on that landscape (and thus are implemented physically as distributed-activation states of the network). To understand why the model works, one must understand how landscape topography and the psychological-to-mathematical implementation relation jointly result in systematically appropriate psychological state-transitions.

For example, in a trained-up network that successfully distinguishes underwater mines from underwater rocks on the basis of input patterns generated from sonar readings, the relevant explanation would take the following form: training the system progressively molds the topography of the network's activation landscape, so that the input patterns generated by rocks each end up situated within an attractor basin whose trajectories all converge on the output representation for rocks, whereas the input patterns generated by mines each end up within the attractor basin of the output representation for mines. In all trained versions of the well-known model NETtalk, the hidden-unit distributed representations of similar letter-to-phoneme correspondences are situated similarly on the activation landscape. Some connectionist models use a sophisticated training regimen called the "moving target strategy," in which input and output representations are modified along with weights. This effects a controlled coevolution of the activation landscape and the psychological-to-mathematical implementation relation – so that these two factors end up "made for each other" with respect to the task the model performs.

In general, connectionist models do not conform to the top-level dynamical hypothesis: they do not posit quantitative psychological variables, or psychological laws expressible as differential (or difference) equations over such variables. However, connectionist models normally do conform to the general dynamical hypothesis; and it is often essential to invoke the mathematical level of description to explain why the model works. Therefore, the dynamical perspective is just as important for connectionist models as it is for models that conform to van Gelder's top-level dynamical hypothesis.

Horgan and Tienson (1996) describe in some detail a nonclassical framework for cognitive science that incorporates what we are here calling the general dynamical hypothesis. We also argue that human cognition is too complex and too supple to conform to psychological laws of the kind posited in classical computational cognitive science – namely, laws that take the form of programmable rules expressible in psychological vocabulary. For essentially the same reasons, we suspect that human cognition is too complex and too supple to conform in general to psychological laws of the kind posited by the top-level dynamical hypothesis. If so, then this is all the more reason for appreciating the alternative possibilities suggested by the general dynamical hypothesis – and for resisting the tyranny of van Gelder's terminology.

## Today's dynamical systems are too simple

Herbert Jaeger

*German National Research Center for Information Technology (GMD), D-53754 Sankt Augustin, Germany;* **herbert.jaeger@gmd.de**
**www.gmd.de/People/Herbert.Jaeger/**

**Abstract:** Cognitive systems are wilder than today's dynamical systems theory can handle. Cognitive systems might be tamed in principle, but the very notion of a dynamical system will change in the process.

Van Gelder stresses the power of recent mathematical insights into dynamical systems. Today we have profound insight into the qualitative phenomenology (attractors, bifurcations, stability) of low-dimensional, input-free systems; we have a working familiarity with chaos, know something about high-dimensional, collective dynamics, and have inklings about spatial and nonstationary systems.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Dynamical systems-oriented investigations into cognitive systems make use of what is mathematically possible today. Typically, a cognitive system is measured or simulated in an approximately stationary mode, that is, with input that most of the time changes much more slowly than on the system's own time-scale. The time series thus obtained are then either *described* as resulting from a certain class of dynamical laws (like in the Decision Field Theory model cited by van Gelder), or they are mathematically *explained* as resulting from universal laws concerning attractors, bifurcations, and chaos-related phenomena (as in the Haken-Kelso-Buny model reported by van Gelder).

The systems studied in this way are typically anatomical or functional subsystems of complete agents. When complete agents are investigated (e.g., Beer 1995b: walking insects; Smithers 1995: maneuvering robots) only a minute fraction of their behavioral repertoire is ever assessed.

Complete cognitive systems arguably have the following properties: (1) they are driven by stochastic input that varies on the system's own characteristic timescales, (2) they are high-dimensional, with many variables developing according to different laws in different subsystems, and (3) and they are nonstationary. Nonstationarity is here understood as resulting from a nonparametric change in the dynamical law itself, as occasioned, for example, by topological restructuring of neural connectivity, by evolutionary processes, or by growth. I will call systems having properties (1)–(3), *wild* systems.

Wild systems cannot be captured by today's mathematicians. Even basic concepts such as attractors or bifurcations cease to be of much help in systems driven by fast, stochastic input. Currently high-dimensional systems can be approached only with respect to some kinds of collective parameters, as in synergetic systems or mean-field approaches. Finally, with nonstationarity of the strong kind noted in (3), we are simply lost.

In my view, cognition can be understood only as a property of complete, situated agents. This is also the view of the interdisciplinary strand of research variously referred to as "situated action," "behavior-based robotics," or "new artificial intelligence" (Pfeifer & Scheier 1998).

Hence, given the premises (1) cognition is irreducibly a property of complete, situated agents; (2) complete, situated agents are wild systems; and (3) current dynamical systems theory cannot catch wild systems, it follows that current dynamical systems theory cannot capture cognitive systems.

One remedy is to bring in the computational hypothesis again, and pursue "hybrid" models of cognition. Examples abound, especially in mobile robotics. Most of these hybrid architectures amount to an addition of dynamical and computational-symbolical modules, but some aim at a true marriage of the dynamical hypothesis with the computational hypothesis with CH (e.g., Jaeger 1994; Hertzberg et al. 1998; Tsotsos 1997). This leads to richer notions of dynamical systems than those in van Gelder's Table 1.

Another remedy is to further develop dynamical systems theory. Unfortunately, progress is slow, because dynamical systems theory is mostly being developed by pure mathematicians and theoretical physicists who have little interest in wild systems. The only exception I am aware of is Casdagli's (1992) generalization of chaos to input-driven systems (but of course there must be others).

More hope rests on control theorists, who combine a mathematical inclination with a professional interest in wild systems. However, most control theorists prefer to domesticize wild systems by linearization, rather than to embark on wild mathematics. The principal advance afforded by control theory seems to be that a notion of optimal behavior and hence of goal-directedness is integrated into dynamical systems theory. Surely some notion of goal-directedness must be included in any satisfactory model of cognitive systems. Again, van Gelder's Table 1 would have to be expanded.

A recent contribution to hunting wild systems is provided by "observable operator models" (OOMs; Jaeger 1998). OOMs can model any stationary stochastic system driven by stochastic input. They come equipped with a constructive learning algorithm that is faster by orders of magnitude than current state-of-the-art iterative procedures (cf. Bengio 1996). The strength of OOMs arises from a reinterpretation of what a dynamical system is. Classically, temporal development is seen as a *succession of states* in some state space. By contrast, an OOM trajectory is a *succession of operators* (hence, "observable operators"). Intuitively, OOMs model temporal development as actions that bring forth new actions. This neatly allows the modelling of thoughts (concepts, propositions, mental images, associations, etc.) as bringing forth successive thoughts. No mechanism or "law" would be needed besides the thoughts themselves. Again, van Gelder's Table 1 would have to grow.

Our mathematical notions are still too narrow when dynamical systems get as wild as true cognition. I feel most grateful to Tim van Gelder for rowing us out into clear waters, but we still have to learn how to lift sails.

## The dynamics of what?

Fred A. Keijzer, Sacha Ben, and Lex van der Heijden

*Unit of Experimental and Theoretical Psychology, Leiden University, 2300 RB Leiden, The Netherlands;* **keijzer@rulfsw.leidenuniv.nl, bem@rulfsw.leidenuniv.nl, heijden@rulfsw.leidenuniv.nl**

**Abstract:** Van Gelder presents the distinction between dynamical systems and digital computers as the core issue of current developments in cognitive science. We think this distinction is much less important than a reassessment of cognition as a neurally, bodily, and environmentally embedded process. Embedded cognition lines up naturally with dynamical models, but it would also stand if combined with classic computation.

Van Gelder's plea for nonclassical approaches to the study of cognition is highly commendable. However, we are uncomfortable with his stark contrast between a dynamical hypothesis (DH) and a computational hypothesis (CH). This very general contrast shifts attention away from explicitly *cognitive* theorizing, which deals with specific questions concerning the nature of cognitive processes.

Van Gelder centers the debate about new approaches to the study of cognition on the question of whether cognition is better explained computationally or dynamically. In van Gelder's treatment, the DH is staged as an alternative hypothesis about the nature of cognitive agents. In his approach, it becomes a matter of prime importance whether the DH is *really* different from the CH or whether the two are at heart equivalent. We think it does not matter much which way the decision will fall. The motivating force that moves cognitive scientists toward the use of dynamical systems theory derives from new ideas about cognition and intelligence themselves. The dynamical language may be particularly congenial for expressing many of those ideas, but (1) we do not see any principled opposition to digital computation, and (2) the driving force behind the current changes in cognitive science consists of changes in specific ideas about cognition itself, not the DH.

The deep issues that are currently debated concern the nature of cognitive processes. New answers have been proposed to the question "What makes something a cognitive system?" Van Gelder avoids this question: "This paper simply takes an intuitive grasp of the issue for granted" (sect. 4, para. 3). However, it is precisely these intuitions that are currently changing. These changed intuitions about what intelligence amounts to are certainly associated with a strong interest in dynamical modeling, but cognitive science turns dynamical because these formalisms fit in more easily with these independently occurring theoretical developments than a classical computational language does.

The most important general change is that cognition is no longer conceptualized as internal thought-processes, but as an intrinsically embedded process in which an agent achieves intelli-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

gent actions by exploiting the physical characteristics of its own body and its environment. Cognition is redrawn as a notion centered around intelligent behavior – or adaptive organism-environment interactions. The newfound importance of timing problems in cognition is a direct corollary of this organism-environment interaction perspective. For successful behavior, it is not only important to decide what should be done, but also in what time-window certain initiated actions will be successful. What cognition consists of is increasingly answered in terms of a capacity for intelligent behavior rather than abstract thought. All of this is mentioned by van Gelder as supporting the DH, but these developments would still stand if someone would prove that they were perfectly reconcilable with the CH.

The derived importance of the DH becomes particularly obvious when a specific theoretical issue is considered: the use of representations in cognitive science. The classical story is that the outward behavior of cognitive agents results from an internally represented program that initiates and guides the external behavioral events, or the output for short – of course always in conjunction with auxiliary feedback loops and other peripheral mechanisms. Van Gelder classifies this general idea with the CH, whereas the DH would provide different options for the use of representations or even their total abandonment.

We find the current dominant interpretation of representation highly problematical; radical changes are needed to fit this notion in cognitive science. Dynamical modeling, however, fails to provide the necessary leverage to initiate a change in the use of representations in cognitive theorizing (Keijzer & Bem 1996). The DH can be perfectly reconciled with a classical interpretation of representations and the CH can be combined with different kinds of representation. Any change in the notion of representation will derive from changes in our ideas about how representational entities function in cognitive theories (Keijzer, in press). The DH is too general to take the leading role here.

The really important issue for cognitive science right now is to extend our understanding of cognitive systems. Whether we will prefer to call these systems computational or dynamical will be a pragmatic decision, depending on what seems to be the most natural and efficient way to explicate and formalize these ideas. Given the direction of current insights into cognitive processes, it seems possible and even plausible that this will ultimately lead to something like the DH. First, however, it should be made clear what all these dynamics are supposed to be the dynamics of.

## Theories of structure versus theories of change

Melanie Mitchell

*Santa Fe Institute, Santa Fe, NM 87501;* **mm@santafe.edu**
**www.santafe.edu/~mm**

**Abstract:** The dynamics/computation debate recalls a similar debate in the evolutionary biology community concerning the relative primacy of theories of structure versus theories of change. A full account of cognition will require a rapprochement between such theories and will include both computational and dynamical notions. The key to making computation relevant to cognition is not making it analog, but rather understanding how functional information-processing structures can emerge in complex dynamical systems.

Is cognition about change or is it about structure? Van Gelder clearly thinks that change is the essence: he champions dynamical systems theory because "dynamicists are interested, in the first instance, in how things change; states are the medium of change, and have little intrinsic interest. Computationalists, by contrast, focus primarily on states; change is just what takes you from one state to another. . . . Computationalists focus on internal structure." (sects. 4.2.3.1 and 4.2.3.2)

Van Gelder's formulation of this opposition – between dynamics as focused on state change and computation as focused on state internal structure – brings to mind a similar debate that has gone on for years in the evolutionary biology community, and whose resolution will, I believe, be instructive for the dynamics/computation debate in cognitive science. What accounts for the biological phenomena we observe in the world? The predominant explanatory framework has been neo-Darwinism, a theory of change par excellence (inherited random change from one generation to the next leads to adaptation by natural selection). However, some evolutionary theorists have questioned the adequacy of classical neo-Darwinism as either an explanatory or a predictive theory, and argue instead for the primacy of historical contingency (Gould 1989a) or the self-organization of biological structure not due to natural selection (Fontana & Buss 1994; Goodwin 1990; Kauffman 1993). These "historicists" and "structuralists" are the connectionists of the evolutionary biology community – the people questioning the classical orthodoxy.

The selectionist/historicist/structuralist debate was summarized by Seilacher in his triangle of causal determinants of form (Seilacher 1991), and has been discussed at length by Gould (1989b), among others. It is becoming increasingly clear, however, that the stark oppositions posited among these three frameworks are not only false ones, but they are hindering progress in evolutionary theory. The purely structuralist theories do not explain how structures can be significantly changed in evolution, and the purely selectionist theories do not explain what intrinsic driving forces and constraints there are on the formation of biological structures. What is needed is a theory that incorporates both change and structure.[1]

Similarly, in cognitive science we have theories of change and movement ("dynamical" approaches): how robots walk in a stable manner (Beer 1995b), how babies reach and grasp (Thelen & Smith 1994), how people move from a condition of uncertainty to making a decision (Busemeyer & Townsend 1993). These theories, however, do not explain the information-processing content of the states over which change is occurring; they either address tasks that do not require complex information processing or they assume high-level information-related primitives a priori. For example, in Busemeyer and Townsend's Decision Field Theory, described in the target article, information-loaded notions such as "positive and negative consequences," "attention and shift of attention," "preferences," and "motivation" are atomic primitives (Busemeyer & Townsend 1993), and the theory does not attempt to explain how these are implemented or why a particular decision-maker might have one version of them instead of another.

Likewise, in cognitive science we have theories of structure ("computational" approaches) that make statements about the information-processing structure of concepts, representations, and beliefs (e.g., semantic networks, neural networks, schemata, Bayesean belief networks, fuzzy logic, theorem provers). As van Gelder points out, most of these theories assume that information processing consists of the manipulation of explicit, static symbols rather than the autonomous interaction of emergent, active ones (Hofstadter 1985). Such theories typically cannot explain what driving forces and constraints there are on how the system in question can change, what trajectories it can take, and how the high-level symbols can emerge from a lower-level substrate.

Thus, as in evolutionary biology, cognitive science needs rapprochements between theories of change and theories of structure. Attempts at such rapprochements are coming from many sectors, in particular from research on "complexity," in which dynamics, computation, and adaptation are beginning to be viewed in a more unified framework. For example, in our work on emergent computation in cellular automata, my colleagues and I have shown how active representations and functional information processing can emerge from interactions among dynamical systems, an environment, and an adaptive evolutionary process (Crutchfield & Mitchell 1995; Das et al. 1994; 1995). This work is a pre-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

liminary step in understanding how useful computation can be embedded in a complex dynamical system; it is one attempt at (as van Gelder puts it) "dramatically reconceiving how [complex internal structures] might be instantiated" in such a system (sect. 6.9). In the end, van Gelder seems to agree that computational notions – albeit of a nontraditional kind – might be important for cognitive science: "the DH can embrace the idea that cognitive processes are computational" (sect. 6.3). However, unlike van Gelder, I do not believe that it is the digital/analog distinction that is key for making computation relevant for cognition; instead, I think progress will come from understanding how functional information-processing structures can emerge in spatially extended dynamical systems with no central control, no globally accessible memory, and limited communication among components. Computer science is gradually moving in this direction, and I believe that many useful synergies between computation theory and cognitive science will arise in the near future.

Van Gelder's answer to objection 6.7 ("not as cognitive") is that according to the DH "cognitive agents are dynamical systems of quite special kinds" (sect. 6.7, para. 2). I will venture to say that they are dynamical systems in which the states and state trajectories can and must be interpreted in functional, informational, and information-processing terms, and the computational notions will be necessary as well as dynamical notions for constructing a full account.

NOTE

**1.** This formulation of the evolution debates was given to me by evolutionist Daniel McShea (personal communication). McShea's formulation was elaborated by Crutchfield (1994), who proposes a particular computation-theoretic notion of structure ("computational mechanics of nonlinear processes") and a related mechanism for the transformation of structure ("hierarchical machine reconstruction"). Crutchfield suggests that a unified theory of these two processes might be termed "evolutionary mechanics," which he proposes as a general theory of "emergence."

## Nature's subtlety undermines the empirical relevance of both dynamical and computational hypotheses

Gregory R. Mulhauser

*Cognitive Systems Research, Applied Research and Technology, British Telecom Laboratories, Martlesham Heath IP5 3RE, England*

**scarab@info.bt.co.uk www.labs.bt.com/people/mulhaug**

**Abstract:** Technical hitches mar van Gelder's proposed map of the conceptual landscape, particularly with respect to descriptive levels and the trio of instantiation, realisation, and implementation. However, for all the formal quibbles, van Gelder is onto something important; the tension he notes between computationalism and a dynamical alternative threatens to transform the way we conduct cognitive science research.

The relationships of instantiation, realisation, and implementation sit at the heart of van Gelder's explication of systems in general and of the computational and dynamical hypotheses in particular. As presented, however, these relationships do not support the tidy dynamical versus computational carving up of the world that van Gelder's project requires.

A central difficulty is that any finite sequence of empirical measurements can in principle be described mathematically in infinitely many distinct ways (the curve fitting of sect. 6.6). Any given object instantiates not only a great many systems (sect. 3.1), but infinitely many. Likewise for implementation: the weak constraint that higher level variables are somehow constructed (sect. 3.1) from those at lower levels leaves the relationship easy prey to "gerrymandering," or wild ad hoc perversions of the spirit of van Gelder's idea that satisfy the letter of his account in unexpected ways. Such inadequately constrained criteria render van Gelder's critically important claims about causal organisation awkward to evaluate. They also invite inconveniences such as Putnam's (1988, pp. 120–25) proof that every ordinary open physical system trivially implements ("instantiates," for van Gelder) every abstract finite state automaton.

Perhaps van Gelder believes that the fundamental relationships can be buttressed easily to support the required distinctions and, crucially, to deliver the big prize of causal organisation that grounds the project. Optimism here places van Gelder in good contemporary company: Chalmers, too, believes that his own single-paragraph description of implementation (Chalmers 1996, p. 318) captures causal organisation. However, neither author acknowledges that the subtleties of pinning down causal organisation occupy entire books (Kitcher & Salmon 1989; Nagel 1961; Rescher 1970; Skryms 1980), having puzzled philosophers of science at least since the inception of the "covering law" model in Hempel and Oppenheim's (1948) seminal paper. The nuances of the problem trace their heritage back, ironically, to David Hume's own scepticism about "necessary connexions" in section VII of his 1748 *Enquiry Concerning Human Understanding*.

Although appeals to the compactness of a description do offer useful constraints, sufficient to preclude some gerrymandering, the concept is tricky. Van Gelder's account of dynamical systems (sect. 3.3, items 1 and 3), hides another glitch: in the general case, the question of how succinctly something may be expressed is formally undecidable. The subject merits a whole field of its own, called algorithmic information theory (Chaitin 1987). At least one usefully nontrivial account of implementation can be constructed with the help of a somewhat tedious information theoretic analysis (Mulhauser 1998), but the resulting view lends no support to a strict demarcation between dynamics and computation.

These quibbles arise more from general problems in the philosophy of science than from the dynamical hypothesis; in fact, the account of computation van Gelder draws from Haugeland (1985) stands, if anything, on even shakier metaphysical footing. Pathological weaknesses in his appeal to correspondence frequently serve for target practice by philosophers such as Sterelny (1990). The upshot? Neither the dynamical nor the computational hypotheses is empirically meaningful in any strong sense: the underlying concepts are too impoverished to support the necessary taxonomy. (Similarly, "all oceans contain fish" would not be an empirical hypothesis if our concept of "fish" were trivially open to philosophical predation.) More relevant to the direction of empirical work is not how different models are classified, but how they enrich our understanding.

Indeed, the rhetorical elegance of van Gelder's treatment should not obscure the fact that both dynamical and computational descriptions, perhaps at different levels, may be invaluable for understanding the same physical object. Real cognitive models rarely fall into a clean explanatory cascade anyway, neatly subsuming one another in relationships of logical supervenience; more often, they are complementary, "horizontally" related, and largely incommensurable. Therefore, van Gelder's persistence in referring to the highest relevant level of causal organisation seems as odd as referring to the highest relevant integer for doing number theory; any particular level or integer may be too high, too low, or just right, depending on the questions being asked. A privileged set of the "right" questions no more exists in cognitive theory than in number theory.

By way of analogy, consider possible explanations of my Macintosh's behaviour while it displays pictures of balls bouncing across the screen. We could start at the (thoroughly dynamical) quantum level, appealing to properties of semiconductor junctions to explain how changes in the computer's internal components and ultimately its display screen occur so reliably and consistently, or we might pick a software level to explain computationally how digital values are copied to a display buffer and make their way to the screen. Alternatively, the best predictor of displayed ball trajectories is probably some class of dynamical systems describing elastic collisions. And in trying to understand "digital" Macintosh balls



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

bouncing behaviour, we might choose a dynamical model precisely because it is useful in this sense. Likewise, we may find it useful to model certain aspects of cognition computationally. Such models can be valuable even though the Macintosh has no actual balls inside and the human brain has no digital logic gates. Van Gelder's rejection of such a view (sect. 4.2.1) with an enigmatic appeal to "relevant similarity" seems altogether unfounded in light of his subsequent observation, with respect to dynamical models of cognition, that there might be no concrete quantity corresponding to individual variables (sect. 4.2.1).

Still, van Gelder is on to something; his work challenges the way we think about cognition, and a heated battle for mind share is under way. However, I believe the field now risks a scientifically unrewarding polarisation, setting dynamics against computation in a competition driven less by empirical research than by philosophically unclear predilections to favour alternative sets of investigative tools. This is a little like mathematicians battling over the virtues of algebraic versus reverse Polish notation calculators while setting the actual mathematics aside.

Both approaches can and should exist side by side, contributing whatever they may to our understanding of the world. The time when science might decide between them is far off, awaiting both philosophical refinement and a clearer empirical picture of natural cognition.

## Is the dynamical hypothesis falsifiable? On unification in theories of cognition

David C. Noelle

*Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213;* **noelle@cnbc.cmu.edu**
**www.cnbc.cmu.edu/~noelle/**

**Abstract:** The dynamical hypothesis is strong in that, for it to be true, *every* cognitive phenomenon must be best modeled by a dynamical system. Depending on how it is interpreted, however, the hypothesis may be seen as probably false or even unfalsifiable. Strengthening the hypothesis to require unification, or at least coherence, across models in different cognitive domains alleviates this problem.

Van Gelder calls researchers away from the bright stroboscopic streetlight of the classical computational approach to cognition, which illuminates but temporally freezes our view of mind, toward the smoothly shifting shadows of dynamical systems. The dynamical hypothesis is a strong claim. It boldly asserts that the best model of each and every cognitive phenomenon is a dynamical one. This is no mere call for the greater utilization of a bag of mathematical tricks, but a proposition concerning the causal structure of every kind of behavior. Still, in at least one way, the dynamical hypothesis is not strong enough. In the form in which it is presented, it may not even be empirically falsifiable.

As summarized in section 4.3, the dynamical hypothesis could be felled by a single counterexample. If but one "kind of cognitive performance" is best characterized as the processing of a digital computer, the dynamical hypothesis is false. The space of behavioral domains that could be searched for such a counterexample is quite large, however, ranging from coordinated finger motion to decision making, and, as pointed out in section 4.1, each domain may require a separate model.

Some have suggested that the human ability to perform mental computations, such as arithmetic, is just the sort of counterexample needed to topple the dynamical hypothesis. As pointed out in section 6.12, however, mental computation need not be subserved by a process of discrete symbol manipulation. Indeed, dynamic quantitative systems in the form of recurrent connectionist networks are capable of learning to perform algorithmic tasks such as mental arithmetic (Cottrell & Tsung 1993).

One potential counterexample that was not addressed by the target article, however, is that of explicit instruction following. The human ability to systematically enact arbitrary symbolic rules, communicated through language, has been used to argue that some aspect of cognition must have a classical computational architecture (Hadley 1998). Unlike mental computation, the domain of instruction following makes the symbolic structure that guides behavior observable, appearing as the explicit rules communicated as instructions. It has been argued that the human capacity to systematically transform such arbitrary symbolic rules into appropriate behavior is evidence for a mental process akin to digital computation. However, this does not necessarily follow. Quantitative connectionist models of instruction following have begun to appear (Noelle 1997), and these capture not only instruction-following behavior but also the interaction between instruction following and various phenomena of language learning, working memory, and automaticity. For instance, when learning a category structure from both instructions and examples, humans sometimes exhibit an interference effect in which similarity between examples hinders proper instruction following (Allen & Brooks 1991); this effect has been displayed in a connectionist model (Noelle & Cottrell 1996b).

When faced with such a quantitative model of instruction following, however, a supporter of the computational hypothesis may concede that the quantitative model best captures issues of interaction and interference with other cognitive processes, but he could still insist that the basic capability of following instructions is best seen as the result of discrete computations. This "classicist" may reject the relevance of the adjacent phenomena, circumscribing a very specific cognitive domain in which only the systematic enactment of simple instructions is considered and only at a coarse temporal granularity (e.g., about 1–10 seconds [Newell 1990]). Under such constraints, the model that most succinctly captures the highest level of causal structure may very well be symbolic and computational. If such a circumscribed cognitive domain is to count as a "kind of cognitive performance" (sect. 4.3), then the dynamical hypothesis is most probably false. If, on the other hand, the supporters of the dynamical hypothesis are allowed to arbitrarily segment the behavioral terrain into domains of interest, there is little hope for the classicist, because some noncomputational phenomena may be included in each domain. And if the notion of a proper "kind of cognitive performance" is left ambiguous, as it is in the target article, then no amount of empirical data will confirm or falsify the dynamical hypothesis.

One way to rescue this hypothesis is to make it even stronger. Rather than requiring a separate dynamical system for each cognitive domain, there must be a unified model for the whole. At the very least, models in particular behavioral domains must be evaluated not only on their ability to capture phenomena in their domain of interest, but also on how well they generalize or relate to models in adjacent domains. A good cognitive model should cohere with models of related behavioral phenomena (e.g., example-based interference in instruction following), models of function and adaptiveness (e.g., general computational accounts, Blum et al. 1998; optimality accounts, Anderson 1990; evolutionary accounts, Barkow et al. 1992;), and biological models (e.g., neural activity). A good theory of a cognitive capability should not only explain a specific class of data, but it should also be rich in these superempirical virtues (Churchland 1989). By strengthening the dynamical hypothesis, requiring a more unified account of cognition, previous doubts concerning falsifiability are dispelled. The existence of any cognitive phenomenon for which there is a digital computational account that both excels at capturing empirical data and coheres well with supported theories in all adjacent domains (including those processes deemed to be noncomputational) would imply that the dynamical hypothesis is false.

In short, the dynamical hypothesis is a strong claim concerning the causal structure of cognition, but it only makes sense as a scientific hypothesis if cognition is seen as a unified whole. Models of cognitive phenomena need to be evaluated in light of their explanatory power in adjacent behavioral domains, in relation to



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

neuroscience, and in relation to the effective functions – the digital or nondigital computations – that they instantiate.

ACKNOWLEDGMENTS

In the spring of 1997, Dan Clouse and I led a seminar in the Department of Cognitive Science at the University of California, San Diego entitled "Cognition as computation versus mind as motion." Some of the ideas expressed here arose from the discussions that took place during this seminar. Thanks are due to all of the participants.

## Digital computers versus dynamical systems: A conflation of distinctions

Gerard O'Brien

*Department of Philosophy, The University of Adelaide, Adelaide, South Australia 5005, Australia;* **gobrien@arts.adelaide.edu.au**
**chomsky.arts.adelaide.edu.au/Philosophy/gobrien/gobrien.htm**

**Abstract:** The distinction at the heart of van Gelder's target article is one between digital computers and dynamical systems, but this distinction conflates two more fundamental distinctions in cognitive science that should be kept apart. When this conflation is undone, it becomes apparent that the computational hypothesis is not as dominant in contemporary cognitive science as van Gelder contends; nor has the dynamical hypothesis been neglected.

In the grand narrative of cognitive science, once we put aside the many internecine squabbles that have enveloped this discipline during its relatively short history, three major developments are discernible. Its beginnings occurred in the 1960s and early 1970s with the application of ideas derived from digital computers to human cognition, spawning the now appropriately named classical computational theory of mind: the doctrine that cognition is a species of symbol manipulation. Then, in the mid-1980s, the field witnessed its first major shake-up with the advent of connectionism: neurally inspired, parallel distributed processing (PDP) models of cognition, which dramatically reconceived the nature of the brain's representational substrate and hence its cognitive architecture. Now the field is once again being thrown into tumult, this time by a group of theorists who advocate abandoning the concept of internal representation altogether, preferring instead to investigate noncomputational ideas and explanatory schemes.

These three developments can be situated in the context of two debates that lie at the very foundations of cognitive science. The first concerns the necessity of invoking computational processes to explain human cognition. The second focuses on the causal structure of the mechanisms that are responsible.

Viewed from a distance, cognitive science is all about the explanation of intelligent behavior. However, intelligence is a very rare commodity in the universe; it is a property lacking in the vast majority of physical systems that inhabit it. One suggestion, in fact one of the founding ideas of cognitive science, is that to explain this aberration it is necessary to invoke a special kind of causal process – the process of computation. Generically speaking, a computation is an operation that processes representations (information-bearing states) in a disciplined fashion. The thought is that intelligent systems, such as ourselves, differ from almost all the other objects in the universe because their behavior is shaped by internal representations that carry information about the environment in which they are embedded. This idea is at the heart of the classical computational theory of mind. It is also the basis of connectionism, although connectionists have a very different view of the representational vehicles implicated in cognition, and hence the kinds of computational processes involved. However, it is precisely this idea that is now being called into question by those at the forefront of the latest development in cognitive science. According to these theorists, the thought that cognition is computational – that it must implicate internal representations of the domain in which an intelligent system is embedded – is actually a hindrance to further development, and ought to be abandoned (see, e.g., Beer 1995; Brooks 1991a).

The picture that emerges when we consider these three developments from the perspective of the second debate is quite different. Although classicists and connectionists are committed to a broadly computational conception of cognition, they disagree about the causal structure of the computational mechanisms the brain uses. Classicists, taking their inspiration from conventional digital computers, view the brain as implementing a coalition of digital devices, and hence hold that cognition is the syntactically governed transformation of symbol structures. Connectionists, by contrast, take their inspiration from the actual hardware found in the brain, and assert that cognition is the result of activation-passing operations across myriad nondigital, PDP-style devices implemented in its neural networks. Along the dimensions of this second issue, therefore, connectionists find themselves in accord with the noncomputationalists, because the latter also champion nondigital causal mechanisms, which they illuminate using the mathematical tools of dynamical systems theory.

It is in this setting that van Gelder articulates the dynamical hypothesis in cognitive science (DH) and defends it as a robust alternative to the computational hypothesis (CH). It is his contention that the latter has dominated cognitive science at the expense of the former, which has been "starved of attention" (sect. 3.3, para. 1). This gives the initial impression that van Gelder's purpose in the target article is to address the first of the debates just reported, and, more important, to formulate a manifesto around which those theorists in the vanguard of the latest (noncomputationalist) development in cognitive science can rally. However, this first impression is quite mistaken. The distinction on which van Gelder's discussion is based, that between digital computers and dynamical systems, actually cuts across the two debates just described, and in doing so conflates two fundamental distinctions that ought to be kept apart. This creates a misleading theoretical geography, as well as an awkward nomenclature. When the appropriate distinctions are more carefully drawn, it becomes apparent that van Gelder's CH is not as dominant in cognitive science as he claims, and that his DH has been far from neglected.

CH in the target article, its name notwithstanding, is not the doctrine that human cognitive processes are computational; it is the narrower claim that these processes are operations of a specific kind of computational mechanism – namely, of digital computers implemented in the brain's neural substrate. Similarly, DH, which is opposed to CH, is not the doctrine that cognitive processes are noncomputational; it is the claim that human cognition is the work of nondigital (i.e., dynamical) causal mechanisms (computational or otherwise). In this light, far from being "starved of attention," DH, in the form of connectionism,[1] has been the subject of intense scrutiny. Moreover, the difference between CH and DH, at least insofar as this revolves around the causal structure of the mechanisms involved in cognition, is what has been at issue in the dispute between classicists and connectionists since the mid-1980s – a debate, it is fair to say, that has dominated cognitive science during this time.

What is new to cognitive science, as already noted, is not an enthusiasm for dynamical models of cognition as such, but a commitment among some of their proponents to an antirepresentational, and hence noncomputational, characterization of the same. This does represent a genuine alternative to both classicism and connectionism, but the debate here is actually obscured by van Gelder's rendering of the theoretical geography. In his terms, this debate pits some advocates of DH against one another, as well as against those who defend CH.

The problem here for van Gelder is that his distinction between CH and DH, based as it is on the distinction between digital computers and dynamical systems, is not fine grained enough to do justice to the current state of cognitive science. What is preferable is a framework that allows for a distinction between digital and dynamical systems on the one hand, and one between computational


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

and noncomputational systems on the other. Only then can the differences between the various theoretical positions that currently occupy the space of cognitive science be adequately characterized.

NOTE

**1.** Witness, for example, Smolensky's (1988) dynamical characterization of connectionism in one of the seminal works on the connectionist program in cognitive science.

## Dynamical modeling and morphological analysis

Jean Petitot

*Mathematical Center, École des Hautes Études en Sciences Sociales, 75006 Paris, France;* **petitot@poly.polytechnique.fr**
**www.ehess.fr/centres/cams/person/petitot.html**

**Abstract:** After a historical sketch of the dynamical hypothesis, we stress that it is a functionalist hypothesis. We then tackle the point of a dynamical approach to constituent structures and emphasize that dynamical modeling must be coupled with morphological analysis.

***The history of the dynamical hypothesis.*** Van Gelder is right that the dynamical hypothesis (DH) in cognitive science has gradually become the most empirically plausible, theoretically relevant, and mathematically interesting one.

As far as I know, it was in the late 1960s that René Thom and Christopher Zeeman (1965) introduced the seminal idea that mental states could be modeled by attractors of (neural) dynamical systems, the temporal flux of consciousness by a "slow" temporal evolution of these "fast" brain dynamics and mental events by the resulting bifurcations of the attractors. They also introduced the key idea that, using order parameters (mean activities), it was possible to reduce the tremendous number of degrees of freedom of such systems drastically and to shift from a micro- (neural) level to a macro- (psychological) level.

These early dynamical models were used in the 1970s by some "structuralist" scientists for modeling perceptual phenomena (such as categorical perception in phonetics or image segmentation in vision) and syntactic structures (in particular semantic roles in case grammar structures). For a summary of these results, see Petitot (1995).

Parallel to this dynamical approach, the connectionist wave (from parallel distributed processing to Smolensky's [1988] works) showed how many cognitive processes could be implemented in explicit neural nets. Now we observe a unification of these two trends in an integrated cognitive dynamical theory.

***Dynamical functionalism.*** Van Gelder strongly emphasizes a key aspect of the DH. The claim that cognitive agents "are" dynamical systems does not concern the dynamical microproperties of the physical substrates these systems are implemented in. It concerns the cognitive level: "the variables [these dynamical models] posit are not low level (e.g., neural firing rates) but, rather, macroscopic quantities at roughly the level of the cognitive performance itself" (sect. 4.1, para. 4). This means that the DH in fact develops a dynamical functionalism. According to functionalism, it is possible to separate the hardware and the software levels: mental states and the relevant relations between them are in some strong sense independent of their concrete implementation. In general, the software levels are conceived of as symbolic, but they can be also dynamical. Indeed, many deep mathematical results on the qualitative theory of dynamical systems (from Poincaré and Birkhoff to Smale, Thom, Arnold, Guckenheimer, etc.) have shown that the behavior of attractors and bifurcations of such systems is highly independent of their microstructure. Hence, there exist emerging dynamical macrostructures and properties that constitute an autonomous level. The main advantage of such a dynamical functionalism is that emergence results from a self-organizing process that causally links the lower microlevels to the higher macrolevels.

***The morphological and structural content of the DH.*** Van Gelder provides the arguments supporting the DH, but, in the end, DH's strength will be measured by its ability to model wide ranges of cognitive phenomena. Some phenomena are very easy to model dynamically, for example, those concerning categorization or learning. Van Gelder comments on other convincing examples, but it must be stressed that one fundamental problem is still wide open. It concerns the capacity of dynamical modeling to provide adequate models for structures: mereological (whole/part) structures, constituent structures, and syntactic structures. Indeed, as has been stressed by Fodor and Pylyshyn (1988) in their critiques against Smolensky's (1988) connectionist approach, high-level cognitive processes require combinatorial syntactic and semantic structures. Only such structures would be able to ensure four fundamental properties of cognitive systems: generativity, systematicity, compositionality, and inferential coherence.

The main advantage of symbolic classical architecture is that these dimensions of systematicity are in some sense built in from the outset. This is not the case with the dynamical approach, which is accordingly obliged to do the job in its own way. This is rather difficult and may require not only dynamical systems but more general models from differential geometry (Petitot 1994b). Let us consider a very simple example: many vision specialists (Marr 1982, etc.) have emphasized that there is a natural constituency or mereology of shapes, namely, a canonical decomposition of every shape in elementary generalized cylinders. This is a key example of constituent structures in visual perception. Now, the best way to generate such a decomposition is to take what is called the skeleton of the shape. Indeed, the skeleton is a one-dimensional graph; its edges are the symmetry axes of generalized cylinders and its vertices indicate how these cylinders must be combined (Blum 1973; Leyton 1992). To get such a morphological analysis of the shape, one makes its boundary trigger a diffusion process: the skeleton is the singular locus of the diffusion. Now, diffusion is driven by a partial differential equation, that is, by a dynamical system in a functional (infinite dimensional) space and not by an ordinary differential equation.

This simple example shows that a dynamical approach must enlarge its mathematical horizon if it aspires to explain even trivial constituent structures. It can be extended widely. Using the "gestalt" approaches to language developed by "cognitive" linguists such as Len Talmy (1983), Ron Langacker (1987), or Terry Regier (1995), it is possible to work out a morphological concept of the constituent structures underlying the systematicity of language (Petitot 1994a). In fact, every structure is a morphology in an appropriate phase space; what we need therefore is a dynamical theory of morphologies in a very general sense, that is a "morphodynamics."

Concerning mathematical tools, classical dynamical systems (with the different meanings analyzed by van Gelder) are only a part of what is required to achieve DH's program. We need not only attractors and bifurcations but also other tools from differential geometry: partial differential equations, singularities, and so on.

## Distinguishing between the computational and dynamical hypotheses: What difference makes the difference?

Steven R. Quartz

*Computational Neurobiology Laboratory, The Salk Institute, La Jolla, CA 92037-1002;* **steve@salk.edu**

**Abstract:** Van Gelder seeks to distinguish between the computational and the dynamical hypotheses primarily on the basis of ontic criteria – the kind of systems cognitive agents really are. I suggest that this meets with mixed



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

success. By shifting to epistemic criteria – what kind of explanations we require to understand cognitive agents – I suggest there is an easier and more intuitive way to distinguish between these two competing views of cognitive agents.

Jerry Fodor once remarked that many of our problems stem from not making enough distinctions. Tim van Gelder is therefore to be congratulated for providing us with much needed distinctions between two alternative visions of cognitive agents, the computational (CH) versus the dynamical hypothesis (DH). Having scoured van Gelder's many distinctions, however, it is not clear to me what differences are the crucial ones, or whether van Gelder believes that there is one such cardinal distinction. Because these are advertised as two very different conceptions of cognitive science, having such a robust, critical distinction seems paramount. Alternatively, if the closest we can get is the definition of section 4.3, entitled "The dynamical hypothesis, exposed" – an extremely dense and not very intuitive formulation – then the DH seems too murky and ill-conceived to win many converts.

There is a simple, basic distinction, one that van Gelder comes close to making many times, but that never quite plays the leading role in a paper that may be too subtle for its own aims. In section 4, for example, he invokes the distinction between the nature and the knowledge hypothesis, but does not use it to full effect. This distinction, sprinkled throughout the target article, is similar to one the computationalists often invoked: levels of explanation. A more explicit use of levels of explanation and a shift from ontic distinctions to epistemic ones could have been used to capture the crucial distinction between CH and DH.

Briefly, computationalists invoked three levels of explanation, the level of mechanism ("functional architecture"), the level of symbol structures, and the semantic, or knowledge level (Newell 1980; Pylyshyn 1984). The critical question occupying much of the computationalist literature concerned the appropriate level of explanation to explain the behavior of cognitive agents (e.g., Stich 1984), making this primarily an epistemological rather than an ontological issue.

I think this suggests that the crucial distinction between the computational and dynamical hypothesis is an epistemic one resting on the appropriate level of explanation for understanding cognitive systems. Van Gelder, however, appears to believe that the distinction between CH and DH should rest primarily on ontological grounds, and so spends considerable effort attempting to determine that cognitive agents are dynamical systems. If the crucial distinction is epistemic rather than ontic, then much of van Gelder's discussion in this regard is peripheral. To cite an example, the discussion of section 6.2 about what kind of system a Macintosh calculating taxes is confuses epistemic and ontic issues. Van Gelder's conclusion (sect. 6.2.3, para. 2), that "while there is one clear sense in which the digital computer 'is' some lower level dynamical system, there is also a clear sense in which it 'is' not that system" confuses rather than clarifies. Ontically, it is either a dynamical system or it is not.

Invoking levels of explanation provides a ready escape. In this way, computationalists are not committed to denying that digital computers are dynamical systems. Of course they are – at some level of explanation. The question is, must we invoke the system's behavior as a dynamical system to explain its calculation of taxes? Probably not, particularly if we want to capture the behavior of all systems running the same program (i.e., those that are functionally equivalent). If the system crashes, however, then we might require a dynamical explanation at the level of functional architecture.

The larger question is, what is the appropriate level of explanation to explain the behavior of a system *as a cognitive agent*? For the computationalist, the behavior of cognitive agents can be captured only at the semantic level, particularly if we are to make generalizations about the behavior of the class of cognitive agents. However, computationalists do not stop there. Instead, they invoke the autonomy of levels of explanation and assert that explanation at the semantic level does not require reference to lower levels of explanation (e.g., Pylyshyn 1984, p. xviii). This autonomy thesis, which rests on the hardware/software distinction and the functionalist views that arose out of it, makes the CH a strong and specific claim about cognitive agents.

Against this background, I think it is fairly straightforward to introduce the dynamical alternative. The DH is rooted in a denial of the autonomy of levels of explanation. It insists on a mixed-level explanation, integrating both the semantic level and the level of the functional architecture to explain the behavior of cognitive agents. Of course, there are many ways to characterize the functional architecture, and so stipulating that the dynamical framework is the appropriate one is an empirical claim, requiring something like the defense van Gelder mounts. This is different, however, from attempting to distinguish between CH and DH on ontological grounds.

Shifting the crucial distinction between CH and DH from an ontic to an epistemic one shifts a number of other issues. For example, in section 4.2.3.9, van Gelder includes antirepresentationalism within the DH tent. I think van Gelder does this on the basis of common ontological commitments. In rejecting the mixed-level explanation, however, representationalists are eliminativist, and so from an epistemological perspective amount to a very different conception of cognitive agents.

The antirepresentational work van Gelder cites in section 4.2.3.9 tends to involve relatively simple systems. Complex internal representations still appear to be necessary for explaining many aspects of human cognition. In my view, the most interesting dynamicist work aims at creating a mixed-level explanation by understanding how representations are instantiated in dynamic properties of the functional architecture (as attractor states, etc.) and by invoking the dynamical language of phase transitions and so forth to substitute for the syntactically governed transitions of computationalism. To properly elucidate such a mixed-level project would, of course, require an extensive treatment. Nonetheless, by locating the crucial difference in epistemic issues, one arrives at an immediate and crucial difference: computationalists assert that there is an autonomous and sufficient semantic level of explanation. Dynamicists deny this, and instead seek to construct a mixed-level explanation, one that integrates explanation at the representational level and the functional architecture by using the tools and concepts of dynamical systems.

There are other significant differences. For example, computationalists see "cognitive agent" as a natural kind, and thus believe that the semantic level would capture the behavior of the entire class of cognitive agents, silicon or neural. By collapsing matters of implementation into those of representation, dynamicists at least implicitly deny this "cognitive universalism," unless a very strong equivalence between the underlying substrates can be shown, perhaps as in analog VLSI implementations. At the same time, however, by collapsing the autonomy of levels, dynamicists disavow the methodological isolationism of cognitive science engendered by the autonomy thesis. This may be the most important feature of the DH, that it invites collaborations across levels of explanation and so disciplines, making cognitive science a very different kind of enterprise from what it once was.

## The intrinsic temporality of human cognition

Benny Shanon

*Department of Psychology, The Hebrew University of Jerusalem, Jerusalem 91905 Israel;* **msshanon@pluto.mscc.huji.ac.il**

**Abstract:** In conformity with the dynamical perspective advocated by van Gelder, a more psychological approach can highlight the intrinsic temporality of human cognition, revealing the inadequacies of representationalism as a framework for the modeling of mind.



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

In complete agreement with van Gelder's call for a dynamic perspective, one can take a more psychological approach and highlight the intrinsic temporality of human cognition. A phenomenon that is usually not sufficiently appreciated is that cognitive acts take time, not only in the sense that actions require a certain amount of time to be executed. Rather, in their very essence, cognitive performances unfold in time. This observation is central in the works of Bergson (1944; 1950; 1913/1983), Husserl (1964), Heidegger (1962), and Merleau-Ponty (1962; 1964). An illustrative example given by Merleau-Ponty (1962) involves the task of evaluating the quality of silk. To check how smooth a fabric is, one has to pass one's hand along the fibers. It is impossible just to touch the fabric and determine its texture. One has to travel along the fiber; this traveling, like all traveling, takes time. It should be emphasized that the extension of the task in time is not due to the collection and integration of more and more information. Rather, the checking is one, single act that, in its very essence, is extended in time.

Until the advent of connectionism, most modern cognitive psychologists ignored time and temporality. The great exception has been James Gibson and his followers in the school of ecological psychology. As pointed out by Gibson (1960; 1966; 1975; 1979), the stimulation provided by the world is not a set of punctate or momentary stimuli; rather, it consists of sensory arrays in flux. Both the inputs that the cognitive system registers and the output it produces gain their identity and their sense through extending in time.

Consider perception. From the representational-computational perspective, perception should not take any time beyond that required for the sensory stimulation to be registered. The time in question is "technical": things take time just as things have spatial dimensions or weight, but there is nothing intrinsically cognitive about this. By contrast, Gibson argues that perceiving is not the reception of a sensory snapshot in a given unit of time in the manner that a camera makes a record of a scene but a dynamic process that is intrinsically temporal. This is a manifestation of the active interaction of the cognitive agent with the world, which is shaped by the past experience of this agent, its present concerns, and its expectations of the future.

The same goes for remembering. Again, the psychological literature is replete with experiments showing that retrieval takes time, that scanning through memory takes time. However, again, these are only technical specifications. The temporality of memory is different. As argued by Casey (1987) and Shanon (1993), remembering cannot be likened to the picking out of an object from a drawer, and the unfolding in time is part and parcel of the memory process itself. In remembering, one does not retrieve pieces of information; rather, one relives an event. Hence, remembering is itself an event, that is, an activity that evolves in time.

The intrinsic temporality of cognition is most apparent in learning. This has been acknowledged by connectionists (see Smolensky 1986). By contrast, classical representational models cannot account for the seemingly banal fact that learning takes time. (In fact, they cannot account for learning at all; see Fodor 1975.) If learning is the incorporation of information presented to the cognitive system from the outside, why should it take time? Why should the acquisition of language take several years rather than one instant? (Indeed, as far as Chomsky is concerned, language acquisition may be – and ideally, is – instantaneous; see Chomsky 1975).

In contrast to all this, orthodox representational-computational models exhibit a principled atemporality. First, the definition of both semantic representations and the computational operations associated with them is nontemporal: in these definitions, there is no reference to temporal specifications. Second, semantic representations are essentially static: they are instantiated as mental states and are attributed with existence regardless of what is done with them. Third, semantic representations are inert: it is assumed that unless they are operated upon they do not change. Fourth, representationalism assumes a fundamental segregation between structures and processes. This segregation is twofold: conceptually, the definition of representations and that of the computational operations associated with them are regarded as two distinct components of the cognitive machine. Finally, the representational framework is ahistorical in that it assumes one can study the present structure and operation of the cognitive system without looking at its past and at the course of its development. As argued at length by Shanon (1993), all these assumptions are wrong.

Against the foregoing characterization, it may be argued that much of experimental cognitive psychology is based on the measurement of response times. Such mental chronometry assumes that mental operations take time. This, however, is not an intrinsic property of the representational system as such. The computational operations postulated in representational models are operations similar to those encountered in logical calculi or transformational grammars. They are defined in an abstract formal domain in which time is not taken into consideration at all. As far as the representational theory is concerned, no temporal constraint is imposed on these operations. In other words, the operations at hand could have taken any length of time – they could have been faster than they actually happen to be, or they could have been slower. There is nothing in the representational theory itself that either attributes temporality to the computational operations or assigns any specific temporal magnitude to them. That they do, in fact, take any time at all, is a technicality that the representational theory accepts as given but about which it has nothing to say. Indeed, there is nothing in cognitive theory that could constrain computational times in any fashion. Representational models specify the course of information processing (i.e., the operations being applied and their sequential order) but not the particular times that they should require (for related discussion, see Pylyshyn 1979).

Representationalism is atemporal, whereas human cognition is intrinsically temporal. This is a categorical difference. As has been noted by the Greek philosopher Zeno, the atemporal cannot be made temporal. Following Zeno's insights, Bergson (1913/1983) pointed out that real time eludes mathematical treatment. Throughout his writings, Bergson (1944 especially) argues that a system comprising fixed, discrete, well-defined terms cannot characterize time. Specifically, the continuous cannot be reduced to the discrete, the ever-changing to the static, the wholistic to the elementary, the multivalued to the univalued. For an extended application of these views to contemporary cognitive science, the reader is referred to Shanon (1993).

## In time and over time

Tim Smithers

*Faculty of Industrial Engineering, University of Navarra, 20099 Donostia, San Sebastián, Gipuzkoa, Spain.* **tsmithers@ceit.es esiiss.ceit.es**

**Abstract:** Van Gelder's clear distinction between the quantitative nature of dynamical systems and the nonquantitative nature of computational processes provides a firm basis for distinguishing between processes that happen *in* time and processes that happen *over* time. Symbolic reasoning, the presumed basis of intelligent behavior in robots, happens *over* time. However, the movements and actions that robots must make to behave intelligently, happen *in* time. Attempting to connect the two, as classical artificial intelligence and robotics have presumed to be necessary, has produced a tension and an arbitrarily moving interface in the construction of robots. Adopting a robotic version of the dynamical hypothesis offers sound theoretical and scientific justification for those robotics researchers who continue to insist that getting the interaction dynamics of intelligent behavior right is a purely dynamical matter, and never a symbolic computational one.

Van Gelder presents a crisp, concise, and much needed clearing of the conceptual and terminological ground between computational and dynamical approaches to cognition. His efforts are of



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

value to cognitive science and to the closely related fields of artificial intelligence (AI) and robotics.

Very similar issues arise in attempts to investigate intelligent behavior in artificial systems where attempts have also been made to develop and promote a dynamical systems approach to intelligent robotics in recent years (Beer 1997; Smithers 1994b; Steinage & Schöner 1997). In many ways these efforts parallel, although on a smaller scale, the motivations, arguments, and developments presented and discussed by van Gelder. This dynamical systems movement in AI and robotics has also generated debate and discussion that has suffered from a lack of clear distinctions and terminology.

In setting out of the dynamical hypothesis in cognitive science clearly, and contrasting it with the more traditional computational hypothesis, van Gelder provides a useful clarification of the issues for researchers in intelligent robotics, and AI more generally. In particular, his definition of dynamical systems as *quantitative* systems – quantitative in time and in state – provides a motivation for interest in a dynamical approach to understanding the mechanisms underlying intelligent behavior in robots.

Most intelligent behavior of interest, in robots or natural systems, necessarily involves actions in the real world: delivering mail, guiding people through buildings, looking for mines, playing table tennis or football, and almost any other kinds of intelligent behavior we would assign to robots. These actions, in turn, depend on making movements; the right kinds of movements, with the right speeds, and at the right times. Movements of the same form, performed at different speeds or times do not in general produce the same actions, hence they do not give rise to the same behavior. The reverse is sometimes also true. Different movements producing different actions can give rise to the same behavior. The rates at which movements are made strongly influence the effectiveness of behavior. Getting these movements and actions right to produce intelligent behavior in robots is a quantitative matter in which it is the dynamics that count. Any theory of the processes and mechanisms underlying intelligent behavior needs to deal with and account for this aspect of intelligent behavior in the real world.

Another way of putting this is that the movements and actions of intelligent behavior happen *in* time: the movements and actions that give rise to intelligent behavior are quantitatively embedded in time. If we change this embedding by slowing down all the movements and actions, or by speeding everything up, we change the behavior, and it will no longer be of much use. Intelligent behavior thus has an important and essential underlying dynamic. Dynamical systems therefore offer useful methods for describing, analyzing, specifying, and realizing the mechanisms needed to control the actions that give rise to intelligent behavior in robots.

Computational (symbol manipulating) processes are, as van Gelder makes clear, not of the same kind; they are not quantitative processes, although they can be used to implement quantitative calculations. Computation, just like everything that happens in the real world, takes time. However, computation happens *over* time, not *in* time. Changing the rate at which the computation is done makes no difference; it remains the same computation. Of course, how long any particular computation takes may strongly affect how useful it is, but this does not change what is computed. Of course, we can stipulate how long some computation can take, for it to be useful or effective in some context, but this is just an extra implementational constraint. We can, and do enclose computational processes in time so as to make them useful or effective in time-constrained applications, such as real-time control or scheduling, but this too is entirely ancillary to the computation being done. Computational processes cannot be made that speed up or slow down depending simply on the state of the computation. To do that, extra hardware, not software is needed.

This fundamental difference between computational or symbol manipulating processes and the mechanisms and processes needed to make the right movements and actions happen in real robots is the cause of a basic tension that has persisted, largely unacknowledged, in AI and robotics since their earliest days. The generally accepted attitude has always been that it will go away with faster computers; it's just a matter of time. The dogma is that symbolic reasoning is necessary for deciding what actions a robot must perform. The problem is to implement this kind of symbolic reasoning so that it is done fast enough to keep up with all the real-time demands of a robot interacting with the real world.

However, as van Gelder illustrates, generating intelligent behavior is not necessarily a matter of computing the right symbolic reasoning fast enough. It is about getting the dynamics of interaction right. So too in robots: intelligent behavior can be understood as getting the interaction dynamics between the robot and its environment right – a necessarily quantitative and dynamical matter. This basic realization is at the origin of the behavior-based robotics movement in AI, a realization that has been only partially accepted by the more traditional robotics research community in their use of behavior-based concepts and techniques – so-called reactive behaviours – for the so-called low-level behaviors of their robots (Brooks 1991b). It is still widely assumed and claimed, however, that producing intelligent behavior in robots only needs reactive behaviors for the low levels, where there is not enough time to do the symbolic reasoning to work out what the robot should be doing. For the "higher level" behaviors, we still need symbolic representation and reasoning methods, even though we still cannot compute it all fast enough.

With time, we will come to understand more about what "in-time" mechanisms and processes we need to produce the well-orchestrated movements and actions of intelligent behavior in real robots. As a result, the ill-defined and essentially arbitrary interface between "over-time" symbolic reasoning processes and the "in-time" processes needed to produce the actual behavior in robots will gradually be pushed further and further up the higher levels of behavior. So, in time, we will have real robots behaving intelligently in the real world that use no kind of computational symbol-processing methods, except perhaps if they play world class chess or prove logical theorems. However, this kind of rule-based reasoning will be realized by suitable dynamical systems, not by the explicit processing of symbols.

## Toward a synthesis of dynamical systems and classical computation

Frank van der Velde and Marc de Kamps

*Unit of Experimental and Theoretical Psychology, Leiden University, 2333 AK Leiden, The Netherlands;* **vdvelde@rulfsw.leidenuniv.nl**
**www.fsw.leidenuniv.nl/~vdvelde/**

**Abstract:** Cognitive agents are dynamical systems but not quantitative dynamical systems. Quantitative systems are forms of analogue computation, which is physically too unreliable as a basis for cognition. Instead, cognitive agents are dynamical systems that implement discrete forms of computation. Only such a synthesis of discrete computation and dynamical systems can provide the mathematical basis for modeling cognitive behavior.

Van Gelder argues that (1) cognitive agents are dynamical systems and (2) these (cognitive) dynamical systems are quantitative systems. We will argue that (1) is correct, but that, in the case of structurally complex cognitive systems, (2) is not.

A key element in the defense of (2) concerns van Gelder's description of computation. He asserts that classical computation is just a specific kind of computation that results from certain constraints on the processes involved (i.e., that they are "effective" or discrete). So other kinds of computation could result from different constraints. In particular, because dynamical systems as quantitative systems are more related to analogue (or continuous) computation, this yields an interesting new branch of computational theory, because analogue computation is more powerful mathe-


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

matically than discrete computation (for a similar argument, see, e.g., Smolensky 1988).

However, the idea that the formulation of discrete computation was based on rather arbitrary constraints is incorrect. Instead, the choice of discrete computation was motivated by *physical* constraints. This can be illustrated with a quote from Turing: "If we admitted an infinity of states of mind, some of them will be 'arbitrarily close' and will be confused" (cited in Epstein & Carnielli 1989, p. 73). The numbers used in analogue computation will, for the most part, be "arbitrarily close" and will get confused in any physical realization, if such a realization is even possible.

In van Gelder's defense of structurally complex cognitive systems as quantitative systems, the issue of the reliability of these systems is ignored. Because of its importance, we will illustrate this point with an example. Consider the recognition or production of the language that consists of sentences of the form $a^n b^n$ (i.e., *ab, aabb, aaabbb,* etc.) This is a context-free language. Therefore, in classical computation theory its recognition or production requires a machine with a stack memory (pushdown automaton).

Pollack (1991), however, presented a "dynamical recognizer," a quantitative dynamical system of the kind advocated by van Gelder, that can recognize (accept) sentences like $a^n b^n$ without the use of a stack. The states $z_i$ of the system consist of "activation" values in the interval [0,1]. The system starts in the state $z_0 = 1$. For each symbol $a$ in the sentence, it changes its state from $z_i$ into $z_{i+1} = \frac{1}{2} z_i$. Therefore, for the first $a$ in $a^n b^n$ it is in the state $z_1 = \frac{1}{2}$, then $z_2 = \frac{1}{4}$, and so on, until it reaches the state $z_n = \frac{1}{2^n}$ For each symbol $b$ in the sentence, it will then change into the state $z_{i+1} = 2z_i$. If the string of $b$s matches the string of $a$s, the end state $z_{end} = 1$. Therefore, the systems accepts the sentence $a^n b^n$ if $z_{end} = z_0$.

This system can be implemented for increasing values of $n$ as long as the states $\frac{1}{2^n}$ and $\frac{1}{2^{n+1}}$ can be distinguished physically. However, even for low values of $n$, these states will be "arbitrarily" close and will probably get confused, as anticipated by Turing.

The fact that these quantitative systems can be simulated on the computer does not carry much weight in this discussion. Indeed, a computer simulation only begs the question. After all, digital computers can be accurate to 10 or 20 decimal orders of significance precisely because they are based on the principles of discrete computation. Only a true analogue implementation can show the viability of these systems as models for cognition. (For an illustration of the difficulties associated with true analogue implementation, see Elias et al. 1997.)

However, doubts about the viability of quantitative dynamical systems do not entail a need to reject the hypothesis that cognitive agents are dynamical systems (point 1 of van Gelder, mentioned earlier). A very persuasive argument, also stated by van Gelder, is that cognitive agents are engaged in an ongoing interaction with the environment. Speed of information processing is an important factor in this interaction. This entails that computational theory is incomplete as the mathematical basis for modeling cognition (van der Velde 1997). Computation theory is about functions, and functions are static in the sense that only the relation between the input and the output matters. Speed of processing is not relevant for this relation. In contrast, because time is a "natural" parameter in dynamical systems, speed of information processing can (or even must) be modeled with dynamical systems.

However, there is a direct relation between dynamical systems and functions. Dynamical systems can be described in terms of sequences of functions (a "flow," e.g., see Jackson 1991). Computational theory is still important for understanding and modeling cognition when the functions that constitute the flow of a cognitive dynamical system are particular computable functions. In that case, a synthesis between dynamical systems and classical computation is achieved (see van der Velde 1997).

Examples can be found in van der Velde (1995; 1997). One example concerns the production of the sentences $a^n b^n$ with a modular dynamical system that consists of attractor neural networks (ANNs). An ANN is capable of storing and recognizing patterns in the form of attractors (e.g., see Amit 1989). In the production of the sentences $a^n b^n$, patterns (attractors) are used to represent the rules (grammar) and the symbols of the language. Pattern recognition with ANNs is also used to implement a stack memory of variable length.

The implementation of a pushdown automaton with ANNs is of course a toy model. However, it illustrates that the productivity of classical computation can be combined with the dynamical characteristics of dynamical systems such as ANNs. Furthermore, the implementation of rules and symbols with patterns (attractors) shows that computational representations are not necessarily static configurations of symbol tokens.

The limited set of attractors that an ANN can store (Amit 1989) are in fact a nice illustration of a discrete set of states in a dynamical system. This again illustrates that the class of discrete computational systems and the class of dynamical systems are not totally different. Indeed, the mathematical basis for modeling cognitive behavior may be found in the intersection of these two classes.

## An appeal for liberalism, or why van Gelder's notion of a dynamical system is too narrow for cognitive science

Michael Wheeler

*Department of Experimental Psychology, University of Oxford, Oxford, OX1 3UD, United Kingdom.* **michael.wheeler@psy.ox.ac.uk**

**Abstract:** Van Gelder identifies the notion of a dynamical system with that of a quantitative system. According to an alternative view, a dynamical system is a state-determined system. This suggests a more profitable way to understand the roles of computation and dynamics in cognitive explanation.

Van Gelder makes a careful and sustained attempt to reveal how the field of dynamics might supply the theoretical foundations for an alternative cognitive science, one that can compete on the playing field of empirical explanation with its more established computational cousin. It is not part of van Gelder's project to explicate the various concepts and principles of dynamics (there are many textbooks that do that), but much hangs on exactly what he means by the term "dynamical system." This is the question I shall address.

Van Gelder argues that the term "dynamical system" should be unpacked as "quantitative system," where a system is quantitative to the extent that its behaviour depends either on distances between its states or on intrasystemic rates of change (sect. 3.3). Quantitative systems contrast sharply with digital computers, "systems whose behaviors are algorithmically specified finite sequences of basic operations constituting manipulations of representations" (sect. 3.2, para. 3) and whose paradigmatic theoretical manifestations are Turing machines (sect. 6.2.2). Digital computers (understood by way of Turing machines), for example, do not exhibit distances between states. Neither do they embody change in time, over and above that of mere ordering (sect. 6.2.2). Deeply indicative of these distinctions is the fact that the variables of digital computers are symbols, whereas the variables of quantitative systems are, as one might expect, numerical quantities (sect. 3.3). On van Gelder's view, then, the terms "digital computer" and "dynamical system" pick out fundamentally different, nonintersecting classes of system. This produces a tension between any dynamical systems approach to cognitive science and the existing computational orthodoxy.

Van Gelder's appeal to Turing machines might well attract a good deal of suspicion from cognitive scientists who would prefer to endorse a less austere understanding of computation (as, e.g., the performance of any well-defined input-output function). However, the Turing machine account not only places the claim



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

that cognition is computation on a solid theoretical footing, but makes that claim substantial and empirically falsifiable. So, on that issue, I think we should follow van Gelder's lead. By contrast, I believe we should question his notion of a dynamical system. In science as a whole, the concept "dynamical system" is often understood not as "quantitative system," but as "state-determined system." (A system is state-determined if its current state always determines a unique future behaviour.) Thus, we would be in good company if we thought of a dynamical system as any system for which we can provide (1) a finite number of variables that (relative to our explanatory interests) capture the instantaneous state of the system, plus (2) a set of equations describing how those variables change. This liberal definition permits the variables of dynamical systems to be either quantities or symbols. Therefore, computational systems and quantitative systems form distinct subsets of the set of dynamical systems. The dimensions along which computational and quantitative systems differ (see sect. 4.2) accordingly emerge as metrics for judging how computational or quantitative a particular dynamical system is. (For an ingenious analysis of Turing machines as dynamical systems, which seems to support this position, see Giunti 1991.)

Van Gelder in fact agrees that a broader notion of a dynamical system would be "reasonable in the light of some strands of contemporary usage" (sect. 6.2.1). However, for the purposes of cognitive science, he dismisses the idea, in part because he thinks that the narrower notion "facilitates articulation and defense of the [dynamical hypothesis]" (sect. 3.3, para. 7). What this indicates is that the choice here is essentially strategic, turning on the philosophical and scientific payoffs that alternative definitions might bestow (which means, of course, that the dispute is not merely terminological). Given such a criterion, the liberal option has much to recommend it. For example, it helps us to resolve what appears to be a tension in van Gelder's account. Van Gelder claims that "Even the loftiest forms of natural cognition are in fact embedded three times over: in a nervous system, in a body, and in an environment" (sect. 5, para. 6), and the dynamical systems approach, unlike the computational approach, is naturally suited to modelling this embeddedness (sect. 5). However, he also grants that some "lofty" expressions of cognition (e.g., performing arithmetic in the head) might literally be digital computation (sect. 6.12). One might resolve this tension in the following way:

(1) One recognises that cognition displays different degrees of embeddedness in brain, body, and environment in different behavioural contexts. (2) One predicts that the best cognitive-scientific explanations will become progressively less computational (atemporal, symbolic) and progressively more quantitative ("more dynamical") either (a) when the cognitive phenomena under investigation are increasingly action oriented, and thus bound up with real-time interchanges between brain, body, and environment (cf. in-the-head arithmetic and playing soccer), or (b) when explaining cognition is seen to require an understanding of the rich dynamics produced by the complex recurrency, interconnectivity, and internal timing details of biological nervous systems. (For a discussion of the latter point in the context of connectionism, see Wheeler 1994.) (3) One notices that the liberal option (unlike van Gelder's view) allows us to understand this progression using a single, integrated conceptual framework. All cognitive phenomena are, it seems, dynamical, but some are more dynamical than others.

Van Gelder's bold and incisive target article goes a long way toward clearing some murky conceptual waters. However, in the end, it is less compelling than it might have been, because cognitive science would be better served by a notion of a dynamical system that is broader than van Gelder's, one that includes, rather than excludes, digital computers.

ACKNOWLEDGMENTS

This work was supported by a Junior Research Fellowship at Christ Church, Oxford, with additional assistance from the McDonnell-Pew Centre for Cognitive Neuroscience, Oxford.

# Author's Response

## Disentangling dynamics, computation, and cognition

Tim van Gelder

*Department of Philosophy, University of Melbourne, Parkville VIC 3052, Australia.* **tgelder@ariel.unimelb.edu.au**
**www.ariel.its.unimelb.edu.au/~tgelder**

**Abstract:** The nature of the dynamical hypothesis in cognitive science (the DH) is further clarified in responding to various criticisms and objections raised in commentaries. Major topics addressed include the definitions of "dynamical system" and "digital computer"; the DH as Law of Qualitative Structure; the DH as an ontological claim; the multiple-realizability of dynamical models; the level at which the DH is formulated; the nature of dynamics; the role of representations in dynamical cognitive science; the falsifiability of the DH; the extent to which the DH is open; the role of temporal and implementation considerations; and the novelty or importance of the DH. The basic formulation and defense of the DH in the target article survives intact, though some refinements are recommended.

The central mission of the target article was to articulate the dynamical hypothesis (DH) and to defend it as an open empirical hypothesis. Much of the article was given over to countering objections to this project, but it would have been impossible to anticipate all reasonable concerns. This response provides an excellent opportunity to continue the effort. It focuses on critical themes common to a number of commentaries, or more likely to arise in the mind of the general reader. They have been classified into three main categories: objections to the target article's characterization of dynamical systems and digital computers; objections to the formulation of the DH; and objections to the defense of the DH as an open empirical hypothesis.

The DH seems to attract the interest of a broad range of cognitive scientists – from roboticists such as **Smithers** to social psychologists such as **Eiser.** Although this response deals mostly with objections or criticisms, it is worth noting that the majority of commentators clearly have a favourable attitude toward the DH; even the most critical concede that there is at least some truth to it. This supports one of the major theses of the target article, that the DH is a serious candidate for the fundamental "Law of Qualitative Structure" concerning human cognition; as such, it is deserving of the sustained empirical investigation it is now receiving in so many parts of cognitive science.

### R1. Objections to the definitions of "dynamical system" and "digital computer"

**Quartz** wonders whether there is any single crucial difference between the dynamical and computational hypotheses. The general answer is no; both are webs of ideas and commitments, and their contrast is multifaceted. Nevertheless, some distinctions are more central than others, and one of the most central is that between dynamical systems and digital computers.

#### R1.1. Is the definition of "dynamical system" too narrow?

The target article's definition of dynamical systems as quan-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

titative systems is criticised for being both too narrow and too broad. **Wheeler** takes the former position; he argues that it would be more profitable to equate dynamical systems with the broader class of state-determined systems. There is much to be said in favour of this approach, some of which he lays out. On the other hand, there are also good reasons for going the "quantitative" route; these are laid out in section 3.3. Wheeler rightly notes that "the choice here is essentially strategic, turning on the philosophical and scientific payoffs." He does not, however, demonstrate that his approach is the more profitable one. That would be difficult, as there is a serious problem lurking in his proposal. The concept of a state-determined system cannot be the right one, because many models with excellent dynamical credentials are not deterministic (e.g., the DFT model, sect. 2). We should not equate dynamical systems with systems in general, however, for that would rob the term "dynamical" of any meaning, and render the DH trivially true. What then is this broader class of systems?

**R1.2. Is the definition of "dynamical system" too broad?** By contrast, **Beer** and **Chater & Hahn** claim that the target article's definition is "excessively general." In different ways, they argue that the category of quantitative systems embraces digital computers such as Turing machines, so the desired contrast between the dynamical systems and digital computers has been lost.

**Beer** claims that digital computers automatically count as quantitative systems because "the integers over which digital computers operate are nothing if not quantitative, and metrics can certainly be defined on integer spaces." However, the state space of the system and the integers are quite different things. Consider a Turing machine computing some function "over" the integers. The Turing machine, considered a system in the sense specified by the target article, is a set of variables; these variables are head state, head position, and locations on the tape. (These are the things that change over time in the operation of the machine.) The state space of the Turing machine is the set of all possible combinations of values of this set of variables. Ontologically, this is wholly different from the integers. Consequently, the fact that the integers constitute a metric space implies nothing about the state space of the Turing machine.

As pointed out in the target article, the state space of the Turing machine is, in fact, a metric space, for certain trivial metrics always apply. However, the target article also required system behaviour to be systematically related to distances as measured by the metric. To this, **Beer** responds: "How can you tell whether a system is bouncing around its state space because its dynamics are chaotic or because its behavior is based on nonmetrical formal properties?" But "how you tell" is irrelevant to the ontological issue (though it may sometimes be a practical concern). By analogy, genuine banknotes are ones produced by the official mint. That is an ontological point. Good counterfeiters can produce notes that are, for all practical purposes, indistinguishable from the real thing. That is an epistemic point. The epistemic problem is not a reason for rejecting the ontological distinction.

**Chater & Hahn**'s case turns on a clever trick. They define the following measure on the states of a Turing machine (or any digital computer): "The distance between two states is the minimal number of steps between them," where "steps" refers to operations of the machine. They then claim that this measure is a metric, and that system behaviour is systematically related to the distances on this measure.

It is hardly obvious that this measure is always a metric; this kind of claim ought to be proved rather than just asserted. But even if it were a metric, there would still be an important difference between paradigmatic examples of dynamical systems and Turing Machines. In the former case we have a generic metric, such as Euclidean distance, defined independently of the behaviour of the system. We can then describe the behaviour of the system using a rule (such as a set of differences or differential equations) expressing behaviour in terms of amounts of change as measured by that behaviour-independent metric. **Chater & Hahn**'s metric, on the other hand, is defined in terms of the behaviour of the system itself. We cannot know what the distances are in the state space until we know how the system behaves. Consequently, we cannot describe system behaviour by specifying amounts of change as measured by a behaviour-independent metric. Put another way, the systematic relation between behaviour and distance as measured by their metric is not a deep and interesting property of Turing machines; it is entirely post hoc.

The upshot is that the target article's definition of "dynamical system" is indeed too broad as it stands, but it needs only a minor correction. That correction is to reformulate the "quantitative in state" condition to require an independent metric, that is, a metric specifiable independent of system behaviour.

**R1.3. Are quantitative systems too analog?** For the purposes of the DH, **van der Velde & de Kamps** argue, dynamical systems should not be defined as quantitative systems. Their reasoning appears to be this: (1) quantitative systems implement analog computation; (2) analog computation cannot be the basis of cognition, because it is too physically unreliable; therefore, (3) cognitive agents cannot be quantitative systems. Yet (4) cognitive agents are dynamical systems; therefore, dynamical systems cannot be quantitative systems.

Their premise (1) needs refinement, but the greatest weakness in their case is premise (2). Is analog computation too unreliable to account for human cognition? Certainly there are implementational issues associated with analog computation, which **van der Velde & de Kamps** illustrate in their discussion of Pollack's (1991) dynamical recognizer. However, they provide no argument that analog computation is too unreliable. On one hand, the reliability of analog computation depends very much on the nature of the computational task and the particular kind of system carrying it out. Even today, some very demanding computational tasks in astronomy are carried out using analog computers. On the other hand, human cognition is itself "unreliable" in certain ways that may be best accounted for using analog dynamical models. For example, Elman (e.g., 1995) has argued that the limitations and errors arising in his connectionist models of language processing are similar to those found in human subjects, which provides evidence in favour of those models.

Digital computation has been the dominant mathematical and conceptual framework for modelling human cognition for some four decades. Numerous cognitive scientists have concluded that, within this approach, they cannot ac-



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

count for the cognitive phenomena of interest, and have been turning to various kinds of dynamical systems, sometimes implementing analog computation. This is not because dynamical systems offer more computational power in some abstract theoretical sense. It is a mathematical truth that certain dynamical systems, construed as analog computers, have super-Turing capacities. However, no dynamicist in cognitive science (to my knowledge) has attempted to make any practical use of such powers. Nobody has taken up dynamical modelling on the promise of super-Turing capacities. Rather, analog dynamical models are pursued precisely because researchers believe that the strengths and weaknesses of real human cognition can best be accounted for by real, physically implementable dynamical models. **Van der Velde & de Kamps** need to provide a more persuasive argument if they want to dismiss all this as misguided.

**R1.4. Are digital computers best defined as interpreted formal systems?** For the purposes of cognitive science, the target article took over the now-standard definition of digital computers as interpreted automatic formal systems (sect. 3.2). It seems that this strategy is acceptable to all commentators except **Chrisley, Mulhauser, Beer,** and perhaps **Bundy.**

**Chrisley** suggests that, for unspecified reasons, the "formal" characterization of digital computation cannot account for computation "in the wild." Similarly, **Mulhauser** claims that it has "pathological weaknesses." Yet neither points to a worked-out, preferable characterization; consequently, notwithstanding their qualms, the response to the "straw man" objection (sect. 6.5) still seems appropriate.

**Beer**'s criticism is that the definition of digital computers is too broad to distinguish dynamical systems and digital computers adequately, because "dynamical systems can easily be defined over discrete state spaces (e.g., cellular automata), and computational descriptions of the symbol sequences generated by discretizing the output of continuous dynamical systems can be used to characterise their complex dynamical structure." This objection seems to be based on a reading of the target article as defining dynamical systems as continuous. This is a misconception (one that **Beer** shares with **Eliasmith** and **Horgan & Tienson**). The target article defines dynamical systems as quantitative; it makes no mention of continuity or discreteness. As Beer rightly notes, quantitative systems can be discrete. Some exemplars of dynamical models in cognitive science are discrete (e.g., van Geert 1995), and so one significant advantage of defining dynamical systems as quantitative is precisely that doing so includes such systems. However, this discreteness does not automatically make the system a digital computer. To be a digital computer, it must satisfy the relevant conditions (sect. 3.2) and discreteness alone is not sufficient.

The fact that sequences of discretized states of continuous dynamical systems can be given (digital) computational descriptions is certainly interesting, but all it really shows is that we can set up complicated mappings between the realms of dynamics and digital computation. It does not show that the dynamical system is a digital computer, any more than the fact that we can simulate the solar system on a digital computer shows that the solar system is a digital computer (see sect. 6.4).

**Bundy** castigates the target article for failing to mention the concept of a virtual machine, an omission he believes seriously compromises its definition of digital computers. However, all the work that can be done with the aid of that concept can be done equally within the target article's framework of instantiation, implementation, and so forth. Introducing yet another term, one that is sure to be interpreted in various ways by different readers, does not necessarily advance the discussion.

## R2. Objections to the formulation of the DH

**R2.1. The DH is not a law of qualitative structure.** The target article presented the DH as a putative law of qualitative structure (LQS). **Braisby et al.** argue that, as formulated, the DH could not be an LQS, because it fails to provide a set of "necessary and sufficient conditions for cognition." Their premise is correct, but their conclusion does not follow, because LQSs are not required to provide necessary and sufficient conditions. The target article borrowed the concept of an LQS from Newell and Simon's (1976) famous Physical Symbol System Hypothesis (PSSH). One of their examples of an LQS is the germ theory of disease, according to which certain diseases are caused by the presence of germs (minute organisms). The germ theory does not provide necessary conditions for disease, because many diseases are not caused by germs, nor does it provide sufficient conditions, because the presence of germs often fails to induce disease (e.g., after inoculation).

In supposing that an LQS must provide necessary and sufficient conditions, **Braisby et al.** have been misled by the form of words Newell and Simon used in formulating the PSSH. Newell and Simon describe the PSSH as the claim that PSSs are necessary and sufficient for general intelligence. Notice, however, that the PSSH does not in fact provide sufficient conditions for intelligence, for not all PSSs are intelligent. In claiming that PSSs are sufficient for intelligence, Newell and Simon meant only that "any PSS of sufficient size can be organized further to exhibit general intelligence." This directly parallels the claim that cognitive agents are dynamical systems of certain quite special kinds (sect. 6.7). If the PSSH counts as an LQS, then so should the DH.

**R2.2. The DH is not a "constitutive account" of cognition.** A related objection is raised by **Chrisley,** who criticizes the DH's failure to present a "constitutive account of cognition," which for him would amount to specifying in dynamical terms necessary and sufficient conditions for something to be cognitive. However, the DH does not purport to provide any such "constitutive account." Its goal is not to tell us what makes something cognitive (sect. 4), but to describe the causal organization of natural cognitive agents. Similarly, the cell doctrine in biology (another Newell & Simon LQS) does not tell us what it is to be alive. Rather, it tells us a contingent fact about life on earth, namely, that the basic building block of living organisms happens to be the cell. Providing a "constitutive account" of cognition, like defining life or justice or piety, may be an interesting philosophical problem, but the DH is a contingent scientific thesis, and none the worse for that.

**R2.3. The DH is epistemological, not ontological.** The DH as articulated in the target article has two major components. The "nature" hypothesis is ontological; it concerns


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

the way the world is. The "knowledge" hypothesis is epistemological; it concerns conceptual frameworks and the proper shape of science. Obviously, the nature hypothesis has priority, because the goal of science is to describe the world ("nature").

Nevertheless, **Quartz, Beer,** and **Bridgeman** all oppose articulating the DH as an ontological hypothesis. The action, they think, really takes place at the epistemological level. According to Beer, for example, "what the debate between computational and dynamical approaches to cognitive science is really about is which is the most insightful, explanatory, penetrating, and parsimonious stance to take toward a cognitive agent." However, none of the three provide a good case for their purely epistemological interpretation.

Both **Quartz** and **Beer** base their positions on perceived difficulties in articulating a clear difference in the ontological commitments of the DH and the CH. We saw above that Beer's objections are misconceived; Quartz's are no better. The bulk of Quartz's discussion of this issue consists of his confessing that he finds the target article's stance on the relations between objects, dynamical systems, digital computers, and so forth a bit confusing. There is certainly a subtle and complex web of relationships here, and keeping them clear in one's mind is not particularly easy. The closest Quartz comes to actually objecting is his flat rejection of the idea that a Macintosh calculating taxes is a dynamical system (in one sense) and is not a dynamical system (in another sense). Against this, Quartz asserts that "in matters of ontology" it is either a dynamical system or it is not." But this is too blunt. The situation (as expressed in the specially tailored vocabulary of the target article) is that the Macintosh, while not itself either a digital computer or a dynamical system, instantiates a low-level dynamical system which implements a digital computer that calculates taxes (see sects. 3.1, 4.1, and 6.2.3 for elaboration of this point).

**Quartz** and **Beer** apparently agree that it makes sense to suppose that there are such things as dynamical systems and digital computers in the world. **Bridgeman** is more radical, rejecting the view that dynamical systems and digital computers are "real world objects." Rather, dynamical or orthodox computational models are just models, not "literal descriptions of the brain," and the choice between them is purely pragmatic.

This kind of antirealism flies in the face of common sense. We generally think of the world as "really" containing systems of various kinds, whose causal structure is the source of the patterns we observe in the world. The planets move as they do because the solar system is perfectly real and behaves the way it does. A calculator produces the answers we need because it instantiates a certain kind of digital computer. When we explain such phenomena scientifically, we produce models purporting to describe literally the systems responsible. Bridgeman would need powerful arguments to overcome this natural attitude, but he provides none. He is right that the brain can be described dynamically or computationally, but this observation is perfectly compatible with a realist interpretation of scientific practice; indeed, the target article sketched such an interpretation.

**R2.4. Dynamical systems are multiply-realizable.** The DH claims that for each kind of cognitive performance we exhibit, there is some dynamical system we instantiate such that those performances are behaviours of that system. Dynamical research in cognitive science aims to identify and understand the system, and it proceeds by proposing abstract systems as models and comparing the behaviour of that abstract system with data on human performance.

**Mulhauser** and **Chrisley** claim that this story conceals a serious problem. There is a famous argument from Putnam (1988) purporting to establish that any abstract formal system is instantiated by every physical object (though the correspondence between the two will almost always be a matter of "wild, ad hoc perversions"). This has the consequence, for example, that any Turing machine can be said to be instantiated by any old bucket of water. Mulhauser and Chrisley suggest that dynamical systems are subject to the same problem: as Chrisley puts it, "one can use Putnam's tricks to show that every physical system instantiates every dynamical system."

It is difficult to assess how urgent this problem is. According to **Chrisley,** the problem in the case of digital computation was solved in 1994 (Chrisley 1994); he worries only that the problem remains for dynamical systems. According to **Mulhauser,** however, the general problem was given a "usefully nontrivial" solution (in Mulhauser 1998). Fortunately, philosophers not only raise problems; they propose solutions as well.

In any case, the truth is that we are all at sea in Neurath's boat, which has some chronic leaks (Quine 1960). In any enquiry we have to treat some problems as provisionally solved (or at least contained) while we address others. Putnam's multiple-realizability objection is such a problem. For one thing, it is not specific to the DH, but is rather a general metaphysical issue. For another, it presents no difficulty for cognitive science in practice. Scientists always judge what counts as a reasonable correspondence between physical object and dynamical system, such that the former is properly said to instantiate the latter. They presume that the sun and planets observe Newton's laws, but do not follow the Hodgkin-Huxley equation, even if some perverse correspondence with the latter could in principle be contrived.

**R2.5. Is the DH pitched at the right level?** The issue of levels, whether in reality or in explanation, is crucial to the formulation of the DH. The nature hypothesis does not say that cognitive agents are dynamical systems at some level; on this reading the DH would be trivial. Rather, it says that cognitive agents are dynamical systems at the level at which we are interested in delivering causal explanations of their behaviour as cognitive agents – or, as the target article put it, "at the highest relevant level of causal organization" (sect. 4.3).

Some commentators question the target article's stance on this issue. **Mulhauser,** for example, doubts there is any single "highest relevant level," and insists that any object might be understood, whether computationally or dynamically, at a number of different levels. The latter point is quite correct, but the former misconstrues the target article's position. The DH does not postulate any single highest level of explanation at which all kinds of cognitive performances are explained. Nevertheless, for any given kind of cognitive performance, there will always be a highest level at which genuine causal explanations of those performances are generated. The highest level might differ, depending on the kind of performance. (Analogously, whereas



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

the highest level for geographic phenomena is that of the tectonic plate, the highest level for photosynthesis is molecular.) The DH says that at each such highest relevant level the explanation will be dynamical, and that at the corresponding level of reality, cognitive agents are dynamical systems.

According to **Quartz,** the essence of the DH is not a substantial thesis about cognitive agents at the highest relevant level. Rather, it is the thesis that natural cognition requires "a mixed level explanation, integrating both the semantic level and the level of the functional architecture." Suppose we call this the "mixed level hypothesis" (MLH); then **Quartz**'s suggestion is that the DH is really the MLH. But the trouble with this idea is that there is nothing intrinsically dynamical about the MLH. Quartz himself opines that a dynamical account of the middle level may be the right way to go, which implies that it might also be the wrong way to go. This makes space for the curious possibility that the DH is true, but that cognitive agents are not modelled as dynamical systems. The MLH may well be right, but it is not the heart of the DH.

**Horgan & Tienson** read the target article as advocating the "top-level" dynamical hypothesis, according to which (a) cognitive agents are best modelled as low-dimensional dynamical systems, and (b) individual variables in the model correspond to top-level "psychological variables" of the agent (e.g., a variable in the model corresponds to the agent's level of preference for some alternative). Horgan & Tienson themselves prefer the "general" DH, which remains neutral on these issues. They regard connectionist models as often falling under the general hypothesis but not the top-level hypothesis.

It is hard to know why **Horgan & Tienson** attribute the top-level hypothesis to the target article. The meagre evidence they provide is outweighed by much other evidence, including the target article's denial that the DH is committed to points (a) and (b) (see e.g., sects. 4.2.1 and 6.8; note that **Eliasmith** makes the same misattributions). Also, the fact that a variety of connectionist models are prominently included among exemplars of the dynamical approach should have suggested that the target article's DH is not restricted to Horgan & Tienson's top-level hypothesis. The DH is in fact much closer to their "general" hypothesis.

**R2.6. Which dynamics?** Part of the knowledge hypothesis is the claim that natural cognition can and should be understood using the theoretical resources of "dynamics." **Crutchfield, French & Thomas, Jaeger,** and **Petitot** all doubt that contemporary dynamics is up to the task. The consensus among the first three seems to be that the dynamical systems underlying much of natural cognition are stochastic, high-dimensional, and nonstationary, and that dynamics as it has been developed to date does not give us much of a grip on these, as Jaeger calls them, "wild" systems. Petitot, on the other hand, emphasises the difficulty of accounting for combinatorial structures within standard dynamics.

The target article was not committed to the idea that dynamics as it exists today suffices for the explanation of cognition; indeed, it noted that dynamics would surely have to be supplemented or even further developed (sect. 6.7). When the knowledge hypothesis refers to "dynamics," it is referring to a suitably expanded future version of dynamics, not the dynamics of the introductory textbooks circa 1998. Of course, on pain of triviality, this future dynamics cannot be simply whatever mathematical framework we eventually find is needed to account for cognition. Rather, it must be recognisably an elaboration of the same generic field of mathematics currently referred to as "dynamics."

**R2.7. The DH is too antirepresentational.** Overwhelmingly, dynamical models in cognitive science incorporate representations of various kinds, purportedly corresponding to representations actually existing in the cognitive system. The target article was attempting to capture the DH as the essence of contemporary dynamical research. Consequently, it characterised the DH as being in opposition to a conception of cognition centered on digital computation, not the use of representation – although it did note that "a small but influential contingent of dynamicists" have in fact rejected representations, at least for their particular purposes (sect. 4.2.3.9).

Nevertheless, some commentators raise the spectre of antirepresentationalism. **Eliasmith** sees the target article as attempting to "all but dismiss the importance of representation to understanding human cognition," and **Dennett** attributes to the "Hard Line van Gelder" the view that "there are no representations at all, anywhere in the brain, in any useful sense – down with representationalism!" Neither cite any evidence of such a view in the target article, and indeed there is none to be found. The most plausible explanation of their mistake is that prior belief in van Gelder's supposed antirepresentationalism biased their interpretation. This only shifts the mystery, however, as nowhere does van Gelder espouse such a view; indeed, in relevant papers he explicitly rejects it (van Gelder 1995, p. 376; van Gelder & Port 1995, pp. 11–12).

## R3. Objections to the defense of the DH

**R3.1. Is the DH falsifiable?** The target article defended the status of the DH as a genuinely empirical hypothesis, one that makes a claim about the way the world is, and one that might well be wrong. This was meant to reflect the experience and practice of dynamical researchers in cognitive science. They produce models of aspects of cognition, and argue for the merits of their models in light of empirical data. They are very much aware that their models, as well as (more generally) their choice of the dynamical framework, must stand or fall according to how well they measure up to the nature of reality as revealed in careful scientific study.

Nevertheless, some commentators express doubts that the DH, as formulated, is an open empirical hypothesis. **Beer** suggests that the knowledge hypothesis "is not a genuine scientific hypothesis, at least not in the traditional sense of making an empirically falsifiable claim. At issue here are not experimentally testable predictions." **Braisby et al.** allege that "imprecision in the statement of the DH renders it unfalsifiable." **Heath** worries that "there is little guidance on how such investigation can determine the relative validity of DH and CH. It may be the case that it is very difficult indeed to provide the empirical evidence needed to reject CH in most cognitive scenarios, using tools available to experimental psychology." And **Noelle** raises the problem that if the notion of "kind of cognitive performance" is up for grabs, partisans will divide up cognition in ways guaranteed to prop up their preferred hypothesis.

About two-thirds of the proper response to each of these


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

objections is proper to all of them. It consists of explaining in what way the DH, like any hypothesis of its kind, is empirically meaningful and hence falsifiable. Obviously, the DH cannot be tested by means of any direct and immediate confrontation with reality. It is a very general hypothesis, perched deep in the web of theory, and surrounded by a wide buffer of auxiliary hypotheses and chains of inference. The DH does, however, issue one major prediction: that our best accounts of cognition will in the long run be dynamical in form. The DH will be known to be false if, after an extensive period of investigation, cognitive scientists have in practice rejected dynamical approaches in favor of some other modelling framework.

In this respect, the DH is on a par with other venerable scientific doctrines. For example, the "evolutionary hypothesis," that all biological complexity is the outcome of natural selection, does not on its own make any specific testable predictions. It does predict however, that in the long run all our best explanations of biological complexity will be cast in terms of natural selection. With much auxiliary theorising, the evolutionary hypothesis does make specific predictions, but if those predictions fail, the main hypothesis can be preserved by shifting the blame elsewhere. If there is too much blame to be shifted, we eventually reject the main hypothesis. For broad theoretical hypotheses, this indirect connection with the world is not nonfalsifiability; rather, it is what falsifiability consists in. Thus, contra **Beer,** the DH can be a genuine scientific hypothesis even if it alone does not make specific testable predictions.

The testability of any broad theoretical hypothesis depends on a fund of good judgement implicit in scientific practice that can never be made fully explicit and written down in a rule book (Kuhn 1962). **Heath** is right to note that in any given case it will be difficult, perhaps impossible, to establish in any conclusive or mechanical way whether a dynamical model is preferable to a computational competitor, but it would be wrong to fault the DH for failing to solve this problem. Appeal to this fund of good judgement also addresses **Noelle**'s concern. At the end of the day we will rely on scientists' intuitive assessments to determine what count as appropriate ways to divide cognitive performances into distinct kinds. If, on one such division, not all cognition is best accounted for in dynamical terms, the DH will have been, to that extent, falsified.

When scientists, as a group, choose one model or general theoretical framework over another, they inevitably allow certain very general desiderata to shape their judgements. Famously, for example, they prefer simple and elegant theories over complex and ungainly rivals. Some refer to such virtues as "aesthetic" or "superempirical;" whatever we call them, it is clear that the process of empirical evaluation always involves relying on such criteria. This is not to say that scientific judgement is "irrational," or "unempirical;" rather, to grasp the essential role of such reliance is to understand the nature of scientific rationality. These desiderata, as **Noelle** reminds us, include comprehensive coverage and smooth integration with theories in other domains.

Finally, is the DH too imprecise to be falsifiable? Perhaps, but the arguments of **Braisby et al.** contain too many errors to establish the point. For example, they claim that the target article provides no criteria for individuating systems. In fact, the target article defines systems as sets of variables, and takes over the standard criterion for set-identity (see Note 5).

**R3.2. Is the DH open?** The target article defended the DH as open, in the sense that whether it is true or false remains to be determined by empirical investigation. More precisely, the challenge is to determine the extent to which the DH is true, because apparently everyone concedes that at least some aspects of cognition are best understood dynamically. The interesting question is whether the DH will turn out to be true of all cognition, including its "highest" or most "central" reaches.

Some commentators are confident that high-level or central cognition is squarely computational rather than dynamical. **Dietrich & Markman** advance what they call a "compatibilist" position: whereas the substrate of cognition as well as some peripheral (sensory and motor) processes are dynamical, the rest is computational. Their case for this position consists partly of confidently repeated assertion, and partly of a quite novel argument. It seems to go like this: Any sort of information-matching or information-comparing process must operate over discrete packets of information. Any process that operates over discrete chunks of information is digital computation. Higher cognition is information processing; therefore, higher cognition is digital computation.

This argument falters at the first step. Why suppose that all information "matching and comparing" processes operate over "discrete chunks" of information? Their only support for this dubious idea is the claim that biologists have found it necessary to regard genes as discrete. Perhaps, but that probably has everything to do with the role of genes and little to do with any supposed "deep point" about information. Moreover, counterexamples are close at hand. Consider using balance scales to make fine discriminations of relative weight. If the weights on each end are analog representations, then we have a nondiscrete information-comparing process. Nondiscrete information processing is found in many kinds of cognitive models. For example, Tuller et al. (1994) unravelled the critical dependence of speech categorisation on subtle (nondiscrete) parameter shifts using an analog dynamical model.

**Bundy** also regards the case against the DH as basically closed when it comes to higher cognition. He offers two kinds of support. One is a report of his take on conventional wisdom ("the consensus"). However, because dynamicists are currently challenging the conventional wisdom, reiterating it just begs the question. His second strategy is more interesting. He describes a particular kind of cognition for which a dynamical account seems "very unlikely," namely, writing a target article for *BBS.* He is right, of course, that it is difficult to say, off the top of one's head, how this process might be given a dynamical account. This illustrates a point made in the target article, that "there are numerous aspects of cognition for which, considered in isolation, the case for dynamical treatment is currently weak at best" (sect. 5). The problem with Bundy's argument is that he fails to mention that accounting for this kind of higher cognitive process is also exceedingly difficult from an orthodox computational perspective. In this case we are on surer ground; we know this, not because our imaginations fail us, but because people have actually tried, and failed dismally, to program digital computers to carry out simpler tasks. Indeed, as a more general point, many people are now inclined to explore the dynamical alternatives, precisely because orthodox computational models, which initially seemed so attractive, turned out to run into very serious problems (Dreyfus 1992).



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

**R3.3. The truth is in the middle.** One possibility is that both the DH and the CH are true, but only in their respective domains. Another possibility is that the whole truth lies in some kind of synthesis of the two. **Crutchfield, Mitchell** and **van der Velde & de Kamps** all recommend an approach centered on understanding how dynamical systems can, under appropriate conditions, give rise to computation or information processing. Mitchell, for example, claims that "progress will come from understanding how functional information-processing structures can emerge in spatially extended dynamical systems with no central control, no globally accessible memory, and limited communication among components."

There is no question that interesting basic research is being done in this area, which **Crutchfield** designates "computational mechanics." However, it is also true that, as cognitive science, this approach is still highly exploratory. It has yet to provide actual models of specific aspects of cognition, accountable to empirical data. When it comes time to model the complexities of real cognition – to publish in *Psychological Review* rather than *Physica D* – they may find that the dynamics drop out of the picture, and the relevant story is cast entirely at the level of the emergent computation. Alternatively, they may find (as have many dynamicists) that the computational aspects play second fiddle to the dynamics. In short, the truth may be in the middle, but at the moment confidence in this position is based more on hunches than on real empirical evidence.

**R3.4. Does the temporality of cognition support the dynamical approach?** One of the major arguments presented by the target article in favor of viewing the DH as a worthy candidate for sustained empirical investigation was the argument from time: briefly, natural cognition unfolds in real time, and dynamical models are better able to account for temporal properties. However, **Chater & Hahn** are unconvinced. They note the interest in issues of time complexity in classical computation theory, and that computationally oriented cognitive scientists have in fact made efforts to deal with timing in their modelling. Where, then, is the advantage of dynamical models?

The target article touched on this area only briefly, but **Shanon** and **Smithers** provide useful elaboration. Shanon argues that many aspects of cognition are intrinsically temporal, and points to a variety of features of traditional computational models supporting the charge of "principled atemporality." Smithers explains the importance of intrinsic temporality to intelligent action from a roboticist perspective. He marks the subtle distinction between the intrinsic temporality of natural cognition and the arbitrary temporality of digital computation with a (perhaps even more subtle) terminological distinction between processes happening "in" time versus those merely happening "over" time. A simple analogy might help here: demerit points accumulate on a driver's license over time, but group jazz improvisation happens in time. In the former case, subtleties of timing are irrelevant; in the latter, they are essential.

Both **Shanon** and **Smithers** do better at describing the inherent temporality of real-world cognition, and the atemporality of digital computation, than at explaining the intrinsic temporality of dynamical systems and the advantage of dynamics in modelling natural cognition. The fundamental point is that in systems exhibiting quantitative state-time interdependence, the time set is not merely an ordered set used to specify the order in which system states are occupied. Rather, it is a metric space, such that amounts of change in state are systematically related to amounts of change in time as measured by that metric. Certain natural systems exhibit this deep "dynamical" property, and so are best modelled with dynamical systems. Digital computers, on the other hand, do not (in general) have this property, and so are weaker models. (See van Gelder, forthcoming, for elaboration.)

**R3.5. Do implementation considerations support the DH?** How do considerations of implementation or embeddedness in bodily (especially neural) hardware bear on the DH? The target article's position was that they constitute one strong prima facie argument in favor of the dynamical approach. **Davids & Bennett** add substance to this case. **Garson** takes a stronger stance: no decent case for the DH can be made without discussing implementation. Because in practice dynamical models are digitally computable, every dynamical model automatically has (Garson claims) a traditional computational competitor exhibiting effectively equivalent behaviour. So how do we distinguish between the two empirically? According to Garson, only by considering how such systems might be implemented in the brain.

There would often be good reason to prefer the dynamical model, however, even if we were completely ignorant of the implementation substrate. When we look closely, the dynamical model and the digital computer simulating it do not exhibit effectively equivalent behaviour. Whereas the dynamical model exhibits the behaviour, the digital computer produces incomplete symbolic descriptions of that model's behaviour. No matter how accurate the description, these are different kinds of things. When I compute the trajectory of a falling body, nothing in fact accelerates at 9.8 $msec^2$. Given this difference, a range of considerations can be brought to bear in determining which of the systems is the more plausible as a model of a particular aspect of cognition, including generic desiderata such as simplicity. It is not the case that only issues of implementation can settle the matter.

A closely related issue is that of the autonomy of cognitive-level dynamics. **Petitot** makes the important point that the DH is what he calls a "functionalist" thesis: macro-level dynamics emerge from lower-level phenomena and have a certain independence from them. A simple but dramatic example might be Feigenbaum's discovery that many dynamical systems undergoing period-doubling bifurcations exhibit a certain ratio property, no matter how they are implemented (Lanford 1982). Dynamicists in cognitive science attempt to describe these emergent, high-level dynamics. An example is the pattern uncovered by Kelso and colleagues, mentioned in section 2: the same fundamental dynamics (captured by variants of the HKB equation) characterise a wide range of phenomena implemented in diverse ways (Kelso 1995). Just as Newton did not need to know anything about the sun and planets except their motions, so dynamicists in cognitive science generally provide models that stand quite independently of implementation details.

Nevertheless, this dynamical functionalism is consistent with **Franceschetti**'s point that there is often "much to be gained" from understanding how macro-dynamics emerge from micro-level structures and processes. We can always go on to ask how the high-level dynamics of cognition are


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

in fact implemented, and there is dynamical research in the process of bridging the gap (see, e.g., Grossberg 1995). As **Bechtel** points out, general explanatory strategies such as mechanistic decomposition may be useful here.

**R3.6. The DH is nothing new.** Dennett once wrote that "there is a familiar trio of reactions by scientists to a purportedly radical hypothesis: (a) 'You must be out of your mind!'; (b) 'What else is new? Everybody knows that!'; and, later – if the hypothesis is still standing – (c) 'Hmm. You might be on to something!' " (Dennett 1995, p. 283). **Dennett** reads the DH as a "purportedly radical" hypothesis, and can not decide whether his response is (a) or (b), and so he allows himself to adopt both simultaneously by projecting onto the target article both an extremist "Hard Line" view and an unremarkable "Soft Line" view. In truth, however, dynamical cognitive science is neither hard-line antirepresentationalism nor just soft reform; it is a genuine science of cognition as different from the orthodox computational approach as soccer is from rugby. Moreover, it is still standing, and so there remains a chance Dennett will eventually adopt the more appropriate attitude, (c).

Variations on reaction (b) – "What else is new?" – are found in the commentaries by **French & Thomas, Eliasmith,** and **O'Brien,** as well. All suggest that the target article presented the dynamical approach as something altogether new, and different from connectionism, whereas it is in fact just good old connectionism in new clothes. They are mistaken in all respects, however. Nowhere did the target article proclaim the dynamical approach as something radically new, and it cited examples of dynamical research going back to the early 1980s. (Van Gelder & Port 1995 provided a historical sketch tracing the dynamical approach back many decades.) In addition, many of the cited examples of dynamical research were connectionist, as were about half the chapters in the dynamical collection *Mind as motion* (Port & van Gelder 1995). Hence the target article did *not* present the DH as, in the words of French & Thomas, "an entirely new modeling paradigm in cognitive science."

It is also important, however, to understand that the dynamical approach is not "just" connectionism. Many dynamical models (including the two examples described in the text, sect. 2) are not connectionist, and connectionists have not always taken a thoroughly dynamical attitude to their models or to cognition. (Anyone supposing that a three-layer feedforward neural net is a good model of cognition is not really a dynamicist, even if the model is in fact a dynamical system.) What really is quite new in all this is just the realisation, in the philosophy of cognitive science, that there is such a thing as a distinctively dynamical approach to the study of cognition, an approach that overlaps with but is not the same as connectionism. This realisation has been around only since the early 1990s, even though connectionist and nonconnectionist dynamical work has been conducted for decades. One major thesis of the target article is that the dynamical/GOFAI divide is more profound and perspicacious than the connectionist/GOFAI divide. Thus if there is any radicalism here, it is not in cognitive science, but one level up, in our thinking about cognitive science.

**O'Brien** objects that the dynamical/GOFAI divide "conflates" the connectionist/GOFAI divide and the representationalist/antirepresentationalist divide. However, he conflates "cross-cuts" with "conflates." From a dynamical point of view, the connectionist/GOFAI opposition is, in O'Brien's words, "not fine-grained enough to do justice to the current state of cognitive science." This claim is elaborated in van Gelder (1997). O'Brien provides no reason to believe that we are better off sticking with the older classifications.

**R3.7. The DH does not matter.** The DH-CH contrast is "much less important" than reconceiving cognition as intrinsically embedded, **Keijzer et al.** claim. Consequently, for them not much turns on the difference between the DH and the CH or which of them we say is true.

It was very much part of the mission of the target article to provide a new conceptual framework (and a revised and regimented vocabulary) for thinking about the fundamental kinds and contrasts in cognitive science. Inevitably, this approach will cut across other ways of thinking about the field, some of which remain important and useful. **Keijzer et al.** observe this tension but do not demonstrate that their preferred alternative is somehow more important or useful than the DH-CH contrast. They do suggest that to discuss the embeddedness of cognitive agents is really to talk about the nature of cognition itself, and is thus somehow prior to talk of the DH-CH contrast. But whether cognitive processes are algorithmic symbolic manipulation or simultaneous coevolution of quantities is just as much a matter of the nature of cognition itself as its embeddedness. Much of the target article, of course, focused on differences among conceptual frameworks, and this discussion necessarily takes place at one remove from the primary subject matter. Nevertheless, the DH-CH dispute itself is a dispute about the nature of cognition.

## References

**Letters "a"' and "r" before author's initials stand for target and response article references, respectively.**


Abraham, E. D., Abraham, R. H. & Shaw, C. D. (1992) Basic principles of dynamical systems. In: *Analysis of dynamic psychological systems, vol. 1: Basic approaches to general systems, dynamic systems and cybernetics,* ed. R. L. Levine & H. E. Fitzgerald. Plenum Press. [aTvG]

Abraham, R. & Shaw, C. D. (1982) *Dynamics - The geometry of behavior.* Aerial Press. [JPC, aTvG]

Ajzen, I. & Fishbein, M. (1980) *Understanding attitudes and predicting social behavior.* Prentice-Hall. [JRE]

Allen, S. W. & Brooks, L. R. (1991) Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General* 120(1):3–19. [DCN]

Allport, G. W. (1935) Attitudes. In: *Handbook of social psychology,* ed. C. Murchison. Clark University Press. [JRE]

Amit, D. J. (1989) *Modeling brain function: The world of attractor neural networks.* Cambridge University Press. [aTvG, FvdV]

Anderson, J. R. (1990) *The adaptive character of thought. Studies in Cognition.* Erlbaum. [DCN]

Anderson, J. R. & Matessa, M. (1997) A production system theory of serial memory. *Psychological Review* 104:728–48. [NC]

Babloyantz, A. & Lourenco, C. (1994) Computation with chaos: A paradigm for cortical activity. *Proceedings of the National Academy of Sciences USA* 91:9027–31. [aTvG]

Baker, G. L. & Gollub, J. P. (1990) *Chaotic dynamics: An introduction.* Cambridge University Press. [DRF, aTvG]

Barkow, J. H., Cosmides, L. & Tooby, J. (1992) *The adapted mind: Evolutionary psychology and the generation of culture.* Oxford University Press. [DCN]

Barnsley, M. (1988) *Fractals everywhere.* Academic Press. [aTvG]

Barton, S. (1994) Chaos, self-organization, and psychology. *American Psychologist* 49(1):5–14. [CE]




https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Bechtel, W. & Richardson, R. C. (1993) *Discovering complexity: Decomposition and localization as strategies in scientific research.* Princeton University Press. [WB]

Beer, R. D. (1995a) Computational and dynamical languages for autonomous agents. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [aTvG]

(1995b) A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence* 72:173–215. [aTvG, HJ, MM, GO]

(1996) Toward the evolution of dynamical neural networks for minimally cognitive behavior. In: *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior,* ed. P. Maes, M. Mataric, J. Meyer, J. Pollack & S. Wilson. MIT Press. [RDB]

(1997) The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems* 20:257–89. [WB, RDB, TS]

Beltrami, E. (1987) *Mathematics for dynamical modeling.* Academic Press. [aTvG]

Bengio, Y. (1996) *Markovian models for sequential data.* TR 1049, Dpt.d'Informatique et Recherche Opérationelle, Université de Montréal. http://www.iro.umontreal. ca/labs/neuro/pointeurs/hmmsTR.ps [HJ]

Bergson, H. L. (1944) *Creative evolution.* The Modern Library. [BS]

(1950) *Time and free will.* Macmillan. [BS]

(1913/1983) *An introduction to metaphysics: The creative mind.* Rowman and Allanheld. [BS]

Bialek, W. & Reike, F. (1992) Reliability and information transmission in spiking neurons. *Trends in Neurosciences* 15(11):428–34. [CE]

Bingham, G. P., Rosenblum, L. D. & Schmidt, R. C. (1997) Dynamics and the orientation of kinematic forms in visual event recognition. *Journal of Experimental Psychology: Human Perception and Performance.* [aTvG]

Blum, H. (1973) Biological shape and visual science. *Journal of Theoretical Biology* 38:205–87. [JP]

Blum, L., Cucker, F., Shub, M. & Smale, S. (1997) *Complexity and real computation: A manifesto.* Springer-Verlag. [aTvG, DCN]

Blum, L., Shub, M. & Smale, S. (1989) On a theory of computation and complexity over the real numbers: NP completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society* 21:1–49. [aTvG]

Brooks, R. (1991a) Intelligence without representations. *Artificial Intelligence* 47:139–59. [GO]

(1991b) Intelligence without reason. In: *Proceedings of the International Joint Conference on Artificial Intelligence, 1991,* Sydney, Australia. [TS]

(1992) Intelligence without representation. In: *Foundations of artificial intelligence,* ed. D. Kirsh. MIT Press. [RLC]

Browman, C. P. & Goldstein, L. (1992) Articulatory phonology: An overview. *Phonetica* 49:155–80. [aTvG]

Bullock, D. & Grossberg, S. (1988) Neural dynamics of planned arm movements: Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review* 95:49–90. [aTvG]

Busemeyer, J. R. & Townsend, J. T. (1993) Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review* 100:432–59. [DRF, aTvG, MM]

(1995) Dynamic representation of decision making. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [aTvG]

Cacioppo, J. T., Gardner, W. L. & Berntson, G. G. (1997) Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review* 1:3–25. [JRE]

Carver, C. S. & Scheier, M. F. (1981) *Attention and self-regulation: A control-theory approach to human behavior.* Springer-Verlag. [JRE]

Casdagli, M. (1992) A dynamical systems approach to modeling input-output systems. In: *Nonlinear modeling and forecasting,* ed. M. Casdagli & S. Eubank. Addison-Wesley. [HJ]

Casey, E. S. (1987) *Remembering: A phenomenological study.* Indiana University Press. [BS]

Casti, J. L. (1992) *Reality rules: Picturing the world in mathematics, vols. 1 and 2.* Wiley. [aTvG]

Chaitin, G. J. (1987) *Information randomness and incompleteness: Papers on algorithmic information theory.* World Scientific. [GRM]

Chalmers, D. J. (1996) *The conscious mind: In search of a fundamental theory.* Oxford University Press. [GRM]

Chomsky, N. (1975) *Reflections on language.* Pantheon Books. [BS]

Chrisley, R. (1994) Why everything doesn't realize every computation. *Minds and Machines* 4:403–20. [RLC, rTvG]

Chrisman, L. (1991) Learning recursive distributed representations for holistic computation. *Connection Science* 3:345–66. [aTvG]

Churchland, P. M. (1989) The ontological status of observables: In praise of the superempirical virtues. In: *A neurocomputational perspective.* MIT Press. [DCN]

Churchland, P. S. & Sejnowski, T. J. (1992) *The computational brain.* Bradford/MIT Press. [CE, aTvG]

Clark, A. (1989) *Microcognition: Philosophy, cognitive science, and parallel distributed processing.* MIT Press. [aTvG]

(1997) *Being there: Putting brain, body and world together again.* MIT Press. [RDB]

Cliff, D., Harvey, I. & Husbands, P. (1993) Explorations in evolutionary robotics. *Adaptive Behavior* 2:73–110. [aTvG]

(1996) Artificial evolution of visual control systems for robots. In: *From living eyes to seeing machines,* ed. M. Srinivasan & S. Venkatesh. Oxford University Press. [WB]

Cohen, A. (1992) The role of heterarchical control in the evolution of central pattern generators. *Brain, Behavior and Evolution* 40:112–24. [aTvG]

Cohen, J. & Stewart, I. (1994) *The collapse of chaos.* Penguin. [aTvG]

Collins, A. M. & Quillian, M. R. (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8:240–47. [RAH]

Copeland, J. (1993) *Artificial intelligence: A philosophical introduction.* Blackwell. [aTvG]

Cottrell, G. W. & Tsung, F.-S. (1993) Learning simple arithmetic procedures. *Connection Science* 5(1):37–58. [DCN]

Crutchfield, J. P. (1994a) The calculi of emergence: Computation, dynamics and induction. *Physica D* 75:11–54. [RDB, JPC]

(1994b) Is anything ever new? Considering emergence. In: *Complexity: Metaphors, models, and reality, Vol. XIX,* ed. G. Cowan, D. Pines & D. Melzner. Santa Fe Institute Studies in the Sciences of Complexity. [JPC, MM]

Crutchfield, J. P. & Mitchell, M. (1995) The evolution of emergent computations. *Proceedings of the National Academy of Sciences, USA* 92:10742. [MM]

Crutchfield, J. P. & Young, K. (1989) Inferring statistical complexity. *Physical Review Letters* 63:105–08. [JPC]

Das, R., Crutchfield, J. P., Mitchell, M. & Hanson, J. E. (1995) Evolving globally synchronized cellular automata. In: *Proceedings of the Sixth International Conference on Genetic Algorithms,* ed. L. J. Eshelman. Morgan Kaaufmann. [MM]

Das, R., Mitchell, M. & Crutchfield, J. P. (1994) A genetic algorithm discovers particle- based computation in cellular automata. In: *Parallel problem solving from nature -PPSN III,* ed. Y. Davidor, H.-P. Schwefel & R. Männer. Springer-Verlag. [MM]

Daugman, J. G. (1990) Brain metaphor and brain theory. In: *Computational neuroscience,* ed. E. L. Schwartz. MIT Press. [KD]

Dennett, D. (1980) The milk of human intentionality. *Behavioral and Brain Sciences* 3:428–30. [RDB]

(1995) *Darwin's dangerous idea.* Touchstone. [rTvG]

Desloge, E. A. (1982) *Classical mechanics.* Wiley. [aTvG]

Dosher, B. A. (1984) Discriminating pre-experimental (semantic) from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology* 16:519–55. [RAH]

Dreyfus, H. L. (1992) *What computers still can't do: A critique of artificial reason.* MIT Press. [arTvG]

Earman, J. (1986) *A primer on determinism.* D. Reidel. [aTvG]

Edelman, G. (1987) *Neural Darwinism.* Basic Books. [KD]

Eiser, J. R. (1994) *Attitudes, chaos, and the connectionist mind.* Blackwell. [JRE]

Eiser, J. R., Claessen, M. J. A. & Loose, J. J. (1998) Attitudes, beliefs and other minds: Shared representations in self-organizing systems. In: *Connectionist and PDP models of social reasoning and social behavior,* ed. S. J. Read & L. C. Miller. Erlbaum. [JRE]

Elias, J. G., Northmore, D. P. M. & Westerman, W. (1997) An analog memory circuit for spiking silicon neurons. *Neural Computation* 9:419–40. [FvdV]

Eliasmith, C. (1996) The third contender: A critical examination of the dynamicist theory of cognition. *Philosophical Psychology* 9(4):441–63. [CE]

(1997) Computation and dynamical models of mind. *Minds and Machines* 7:531–41. [CE]

Elman, J. (1990) Finding structure in time. *Cognitive Science* 14:179–211. [RMF]

(1995) Language as a dynamical system. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [JWG, arTvG]

*Encyclopaedia of Mathematics* (1989). Kluwer. [aTvG]

Epstein, R. L. & Carnielli, W. A. (1989) *Computability: Computable functions, logic, and the foundations of mathematics.* Wadsworth & Brooks/Cole. [FvdV]

Falkenhainer, B. & Forbus, D. F. (1989) The structure-mapping engine: Algorithm and examples. *Artificial Intelligence* 41:1–63. [NC]

Fazio, R. H. (1990) Multiple processes by which attitudes guide behavior. The MODE model as an integrative framework. In: *Advances in experimental social psychology, vol. 23,* ed. M. P. Zanna. Academic Press. [JRE]

Fodor, J. A. (1975) *The language of thought.* Harvard University Press. [aTvG, BS]

Fodor, J. A. & Pylyshyn, Z. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28:3–71. [aTvG, JP] Also in: *Connections and symbols,* ed. S. Pinker & J. Mehler. MIT Press. [ED]



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

Fontana, W. & Buss, L. W. (1994) What would be conserved if 'the tape were played twice'? In: *Complexity: Metaphors, models and reality,* ed. G. Cowan, D. Pines & D. Meltzer. Addison-Wesley. [MM]
Freeman, W. J. & Skarda, C. A. (1990) Representations: Who needs them? In: *Brain organization and memory cells, systems and circuits,* ed. J. L. McGaugh, J. L. Weinberger & G. Lynch. Guildford Press. [aTvG]
Garey, M. R. & Johnson, D. S. (1979) *Computers and intractability: A guide to the theory of NP-completeness.* W. H. Freeman. [NC]
Garson, J. (1996) Cognition poised at the edge of chaos: A complex alternative to a symbolic mind. *Philosophical Psychology* 9:301–21. [aTvG]
Gibson, J. J. (1960) The concept of the stimulus in psychology. *American Psychologist* 16:694–703. [BS]
(1966) The problem of temporal order in stimulation and perception. *Journal of Psychology* 62:141–49. [BS]
(1975) Events are perceivable but time is not. In: *The study of time,* ed. J. T. Fraser & N. Lawrence. Springer-Verlag. [BS]
(1979) *The ecological approach to visual perception.* Houghton Mifflin. [BS]
Giunti, M. (1991) Computers, dynamical systems, phenomena, and the mind. Ph.D. thesis, Department of History and Philosophy of Science, Indiana University. [MW]
(1997) *Computation, dynamics, and cognition.* Oxford University Press. [aTvG]
Globus, G. G. (1992) Toward a noncomputational cognitive neuroscience. *Journal of Cognitive Neuroscience* 4:299–310. [KD]
Goodwin, B. C. (1990) Structuralism in biology. *Science Progress* 74:227. [MM]
Gould, S. J. (1989a) *Wonderful life: The Burgess shale and the nature of history.* W. W. Norton. [MM]
(1989b) A developmental constraint in Cerion, with comments on the definition and interpretation of constraints in evolution. *Evolution* 43(3):516–39. [MM]
Gregson, R. A. M. (1993) Learning in the context of nonlinear psychophysics: The Gamma Zak embedding. *British Journal of Mathematical and Statistical Psychology* 46:31–48. [aTvG]
(1995) *Computation, dynamics, and cognition.* Oxford University Press. [aTvG]
Grossberg, S. (1995) Neural dynamics of motion perception, recognition learning and spatial attention. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [rTvG]
Grossberg, S. & Gutowski, W. E. (1987) Neural dynamics of decision making under risk: Affective balance and cognitive-emotional interactions. *Psychological Review* 94:303–18. [aTvG]
Grossberg, S. & Rudd, M. E. (1992) Cortical dynamics of visual motion perception: Group and element apparent motion. *Psychological Review* 99:78–121. [aTvG]
Grossberg, S. & Stone, G. O. (1986) Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. *Psychological Review* 93:46–74. [aTvG]
Grzegorczyk, A. (1957) On the definitions of computable real continuous functions. *Fundamenta Mathematica* 44:61–71. [aTvG]
Guckenheimer, J., Gueron, S. & Harris-Warrick, R. (1993) The dynamics of a conditionally bursting neuron. *Philosophical Transactions of the Royal Society of London B* 341. [aTvG]
Hadley, R. F. (1998) Connectionism and novel combinations of skills: Implications for cognitive architecture. Technical Report SFU CMPT TR 1998–01, School of Computing Science, Simon Fraser University. [DCM]
Haken, H. & Stadler, M., eds. (1990) *Synergetics of cognition.* Springer-Verlag. [aTvG]
Halford, G. S., Maybery, M. T., O'Hare, A. W. & Grant, P. (1994) The development of memory and processing capacity. *Child Development* 65:1338–56. [RAH]
Harnad, S. (1990) The symbol grounding problem. *Physica D* 42:335–46. [NB, aTvG]
Harvey, I. (1992) Untimed and misrepresented: Connectionism and the computer metaphor. University of Sussex Cognitive Science Research Paper No. 245. [aTvG]
Haugeland, J. (1978) The nature and plausibility of cognitivism. *Behavioral and Brain Sciences* 1:215–26. [aTvG]
(1985) *Artificial intelligence: The very idea.* MIT Press. [DD, aTvG, GRM]
Heft, H. (1989) Affordances and the body: An intentional analysis of Gibson's ecological approach to visual perception. *Journal for the Theory of Social Behaviour* 19:1–30. [KD]
Heidegger, M. (1962) *Being and time.* Harper and Row. [BS]
Hempel, C. G. & Oppenheim, P. (1948) Studies in the logic of explanation. *Philosophy of Science* 15:135–75. [GRM]
Hertzberg, J. et al. (1998) A framework for plan execution in behavior-based robots. In: *Proceedings of the Joint ISIC/ISAS'98 conference on "Intelligent control of complex systems."* http://www.gmd.de/People/Herbert.Jaeger/Publications.html [HJ]
Hirsch, M. (1984) The dynamical systems approach to differential equations. *Bulletin of the American Mathematical Society* 11:1–64. [aTvG]
Hobbes, T. (1651/1962) *Leviathon.* Collier Books. [aTvG]
Hock, H. S., Kelso, J. A. S. & Schöner, G. (1993) Bistability, hysteresis and loss of temporal stability in the perceptual organization of apparent motion. *Journal of Experimental Psychology: Human Perception and Performance* 19:63–80. [aTvG]
Hofstadter, D. R. (1985) Waking up from the Boolean dream, or, subcognition as computation. In: *Metamagical themas.* Basic Books. [MM]
Horgan, T. E. & Tienson, J. (1996) *Connectionism and the philosophy of psychology.* MIT Press. [DRF, aTvG, TH]
Hume, D. (1978) *A treatise of human nature, (1739–40).* Clarendon Press. [aTvG]
Husbands, P., Harvey, I. & Cliff, D. (1995) Circle in the round: State space attractors for evolved sighted robots. *Robotics and Autonomous Systems* 15:83–106. [aTvG]
Husserl, E. (1964) *The phenomenology of internal time-consciousness.* Indiana University Press. [BS]
Jackson, E. A. (1991) *Perspectives of nonlinear dynamics, vols. 1 and 2.* Cambridge University Press. [FvdV]
Jaeger, H. (1994) Dynamic symbol systems. Ph. D. thesis, University of Bielefeld. http://www.gmd.de/People/Herbert.Jaeger/Publications.html [HJ]
(1996) Dynamische Systeme in der Kognitionswissenschaft. *Kognitionswissenschaft* 5:151–74. [aTvG]
(1998) A short introduction to observable operator models for stochastic processes. In: *Proceedings of the Cybernetics and Systems '98 Conference, vol.* 1, ed. R. Trappl, Austrian Society for Cybernetic Studies. http://www.gmd.de/People/ Herbert.Jaeger/Publications.html [HJ]
Jalmy, L. (1983) How language structures space. In: *Spatial orientation: Theory, research and application,* ed. H. Pick & L. Acredolo. Plenum Press. [JP]
Kampis, G. (1991) *Self-modifying systems in biology and cognitive science.* Pergamon Press. [aTvG]
Kaplowitz, S. A. & Fink, E. L. (1992) Dynamics of attitude change. In: *Analysis of dynamic psychological systems, vol. 2: Methods and applications,* ed. R. L. Levine & H. E. Fitzgerald. Plenum Press. [aTvG]
Kauffman, S. A. (1993) *The origins of order: Self-organization and selection in evolution.* Oxford University Press. [RAH, MM]
Keane, M. T., Ledgeway, T. & Duff, S. (1994) Constraints on analogical mapping - A comparison of 3 models. *Cognitive Science* 18:387–43. [NC]
Keijzer, F. A. (in press) Doing without representations which specify what to do. *Philosophical Psychology.* [FAK]
Keijzer, F. A. & Bem, S. (1996) Behavioral systems interpreted as autonomous agents and as coupled dynamical systems: A criticism. *Philosophical Psychology* 9:323–46. [FAK]
Kelso, J. A. S. (1995) *Dynamic patterns: The self-organization of brain and behavior.* MIT Press. [KD, DRF, arTvG]
Kelso, J. A. S., DelColle, J. & Schöner, G. (1990) Action-perception as a pattern formation process. In: *Attention and performance XIII,* ed. M. Jeannerod. Erlbaum. [aTvG]
Kelso, J. A. S., Ding, M. & Schöner, G. (1992) Dynamic pattern formation: A primer. In: *Principles of organization in organisms,* ed. E. Mittenthal & A. B. Baskin. Addison-Wesley. [aTvG]
Kitcher, P. & Salmon, W. C., eds. (1989) Scientific explanation. In: *Minnesota Studies in the Philosophy of Science, vol. 13.* University of Minnesota Press. [GRM]
Krantz, D. H., Luce, R. D., Suppes, P. & Tversky, A. (1971) *Foundations of measurement.* Academic Press. [aTvG]
Kugler, P. N., Kelso, J. A. S. & Turvey, M. T. (1980) On the concept of coordinate structures as dissipative structures: I. Theoretical lines of convergence. In: *Tutorials in motor behavior,* ed. G. E. Stelmach & J. Requin. North Holland. [aTvG]
(1982) On the control and coordination of naturally developing systems. In: *The development of movement control and coordination,* ed. J. A. S. Kelso & J. E. Clark. Wiley. [aTvG]
Kuhn, D. (1962) *The structure of scientific revolutions.* University of Chicago Press. [rTvG]
Lanford, O. E. (1982) A computer-assisted proof of the Feigenbaum conjectures. *Bulletin of the American Mathematical Society* 6:427–34. [rTvG]
Langacker, R. (1987) *Foundations of cognitive grammar.* Stanford University Press. [JP]
Langton, C. G. (1990) Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D* 42:12–37. [RAH]
Leven, S. J. & Levine, D. S. (1996) Multiattribute decision making in context: A dynamic neural network methodology. *Cognitive Science* 20:271–99. [aTvG]
Leyton, M. (1992) *Symmetry, causality, mind.* MIT Press. [JP]
Luenberger, D. G. (1979) *Introduction to dynamic systems: Theory, models and applications.* Wiley. [aTvG]
Maddy, P. (1990) *Realism in mathematics.* Clarendon Press. [aTvG]
Marr, D. (1982) *Vision.* Freeman. [JP]
McClelland, J. L. & Rumelhart, D. E. (1981) An interactive-activation model of



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

context effects in letter perception: Part I, an account of basic findings. *Psychological Review* 88:375–407. [aTvG]
McClelland, J. L., Rumelhart, D. E. & The PDP Research Group, eds. (1986) *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2, Psychological and biological models.* MIT Press. [aTvG]
McCulloch, W. & Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. In: *Embodiments of mind,* ed. W. McCulloch. MIT Press. [BB]
Merleau-Ponty, M. (1962) *Phenomenology of perception.* Routledge and Kegan Paul. [BS]
(1964) *Le visible et l'invisible: Suivi de notes de travail.* Gallimard. [BS]
Mill, J. S. (1975) *On liberty.* Norton. [aTvG]
Moore, C. (1991) Generalized shifts: Unpredictability and undecidability in dynamical systems. *Nonlinearity* 4:199–230. [aTvG]
(1996) *Dynamical recognizers: Real-time language recognition by analog computers. No. 96–05–023.* Santa Fe Institute. [aTvG]
Mpitsos, G. J. (in press) Attractor gradients: Architects of developmental organization. In: *Identified neurons: Twenty-five years of progress,* ed. J. L. Leonard. MIT Press. [aTvG]
Mulhauser, G. R. (1998) *Mind out of matter: Topics in the physical foundations of cognitive science.* Kluwer. [rTvG, GRM]
Nagel, E. (1961) *The structure of science: Problems in the logic of scientific explanation.* Routledge & Kegan Paul. [GRM]
Newell, A. (1980) Physical symbol systems. *Cognitive Science* 4:135–83. [DD, aTvG, SRQ]
(1990) *Unified theories of cognition.* Harvard University Press. [aTvG, DCN]
Newell, A. & Simon, H. (1976) Computer science as empirical enquiry: Symbols and search. *Communications of the Association for Computing Machinery* 19:113–26. [NB, arTvG]
Noelle, D. C. (1997) *A connectionist model of instructed learning.* Ph. D. thesis, University of California, San Diego, Department of Computer Science and Engineering, Department of Cognitive Science. [DCN]
Noelle, D. C. & Cottrell, G. W. (1996a) In search of articulated attractors. In: *Proceedings of the 18th Annual Conference of the Cognitive Science Society,* ed. G. W. Cottrell. Erlbaum. [aTvG]
(1996b) Modeling interference effects in instructed category learning. In: *Proceedings of the 18th Annual Conference of the Cognitive Science Society,* ed. G. W. Cottrell. Erlbaum. [DCN]
Norton, A. (1995) Dynamics: An introduction. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [aTvG]
Nowak, A., Szamrej, J. & Latané, B. (1990) From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review* 97:363–76. [JRE]
Ott, E. (1993) *Chaos in dynamical systems.* Cambridge University Press. [NC, aTvG]
Padulo, L. & Arbib, M. A. (1974) *System theory: A unified state-space approach to continuous and discrete systems.* W. B. Saunders. [aTvG]
Parsegian, V. L. (1973) *This cybernetic world of men, machines and earth systems.* Doubleday. [aTvG]
Petitot, J. (1985a) *Les catastrophes de la parole.* Maloine. [aTvG]
(1985b) *Morphogenèse du sens.* Presses Universitaires de France. [aTvG]
(1994a) *Morphodynamics and cognitive grammar: Continuity in linguistic semantics.* Benjamin. [JP]
(1994b) *Dynamical constituency: Sémiotiques.* Didier. [JP]
(1995) Morphodynamics and attractor syntax. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [ED, aTvG, JP]
Pfeifer, R. & Scheier, C. (1998) *An introduction to new artificial intelligence.* MIT Press. [HJ]
Pollack, J. B. (1990) Recursive distributed representations. *Artificial Intelligence* 46:77–105. [aTvG]
(1991) The induction of dynamical recognizers. *Machine Learning* 7:227–52. [JWG, aTvG, FvdV]
Port, R., Cummins, F. & McAuley, J. D. (1995) Naive time, temporal patterns, and human audition. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. Bradford Books/MIT Press. [aTvG]
Port, R. & van Gelder, T. J. (1995) *Mind as motion: Explorations in the dynamics of cognition.* MIT Press. [RDB, arTvG]
Pour-El, M. & Richards, J. (1989) *Computability in analysis and physics.* Springer-Verlag. [aTvG]
Putnam, H. (1988) *Representation and reality.* MIT Press. [RLC, GRM]
Pylyshyn, Z. W. (1979) Do mental events have durations? *Behavioral and Brain Sciences* 2:277–78. [BS]
(1984) *Computation and cognition: Toward a foundation for cognitive science.* Bradford Books/MIT Press. [aTvG, SRQ]
Quine, W. V. O. (1960) *Word and object.* Harvard University Press. [rTvG]
Rapp, P. E. (1993) Chaos in the neurosciences: Cautionary tales from the frontier. *Biologist* 40:89–94. [RMF]
Ratcliff, R. & McKoon, G. (1982) Speed and accuracy in the processing of false statements about semantic memory. *Journal of Experimental Psychology: Learning, Memory and Cognition* 8:16–32. [RAH]
Read, S. J., Vanman, E. J. & Miller, L. C. (1997) Connectionism, parallel constraint satisfaction processes, and Gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review* 1:26–53. [JRE]
Reed, E. S. (1993) The intention to use an affordance: A conceptual framework for psychology. In: *Development in context: Acting and thinking in specific environments,* ed. R. H. Wozniak & K.W. Fischer. Erlbaum. [KD]
Regier, T. (1995) A model of the human capacity for categorizing spatial relations. *Cognitive Linguistics* 6(1):63–88. [JP]
Reike, F., Warland, D., de Ruyter van Steveninck, R. & Bialek, W. (1997) *Spikes: Exploring the neural code.* MIT Press. [CE]
Rescher, N. (1970) *Scientific explanation.* Free Press. [GRM]
Rosen, R. (1985) *Anticipatory systems.* Pergamon. [aTvG]
Rosenblatt, F. (1960) Perception simulation experiments. *Proceedings of the Institute of Radio Engineers* 48:391–409. [BB]
Rumelhart, D. E., McClelland, J. L. and The PDP Research Group, eds. (1986a) *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1: Foundations.* MIT Press. [aTvG]
Rumelhart, D. E., Smolensky, P., McClelland, J. L. & Hinton, G. E. (1986b) Schemata and sequential thought processes in PDP models. In: *Parallel distributed processing: Explorations in the microstructure of cognition,* ed. J. L. McClelland, D. E. Rumelhart & the PDP Research Group. MIT Press. [aTvG]
Ryle, G. (1984) *The concept of mind (1949).* University of Chicago Press. [aTvG]
Saltzman, E. (1995) Dynamics and coordinate systems in skilled sensorimotor activity. In: *Mind as motion: Explorations in the dynamics of cognition,* R. Port & T. van Gelder. MIT Press. [aTvG]
Saltzman, E. L. & Munhall, K. G. (1989) A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1. [aTvG]
Schmidt, R. C. & Turvey, M. T. (1994) Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics* 70. [aTvG]
Schöner, G., Zanone, P. G. & Kelso, J. A. S. (1992) Learning as a change of coordination dynamics: Theory and experiment. *Journal of Motor Behavior* 24:29–48. [aTvG]
Searle, J. R. (1980) Minds, brains and programs. *Behavioral and Brain Sciences* 3:417–58. [aTvG]
(1990) Is the brain a digital computer? *Proceedings and Addresses of the American Philosophical Association* 64. [RLC]
Seilacher, A. (1991) Self-organizing mechanisms in morphogenesis and evolution. In: *Constructional morphology and evolution,* ed. N. Schmidt-Kittler & K. Vogel. Springer-Verlag. [MM]
Shanon, B. (1993) *The representational and presentational: An essay on cognition and the study of mind.* Harvester-Wheatsheaf. [BS]
Siegelmann, H. T. & Sontag, E. D. (1994) Analog computation via neural networks. *Theoretical Computer Science* 131:331–60. [aTvG]
Skarda, C. A. & Freeman, W. J. (1987) How brains make chaos to make sense of the world. *Behavioral and Brain Sciences* 10:161–95. [aTvG]
Skryms, B. (1980) *Causal necessity.* Yale University Press. [GRM]
Smale, S. (1980) *The mathematics of time: Essays on dynamical systems, economic processes and related topics.* Springer-Verlag. [JPC]
Smith, B. C. (1996) *On the origin of objects.* MIT Press. [RLC, aTvG]
(in preparation) *The age of significance: An essay on the foundations of computation and intentionality.* [aTvG]
Smith, E. R. (1996) What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology* 70:893–912. [JRE]
Smith, L. B. & Thelen, E. (1993) *Dynamic systems in development: Applications.* MIT Press. [aTvG]
Smithers, T. (1994a) On behavior as dissipative structures in agent-environment system interaction spaces. In: *Proceedings of Prerational Intelligence: Phenomenology of Complexity in Systems of Simple Interacting Agents, Zentrum für interdisziplinäre Forschung (ZiF), University of Bielefeld, Germany.* [aTvG]
(1994b) What the dynamics of adaptive behaviour and cognition might look like in agent-environment interaction systems. In: *On the role of dynamics and representation in adaptive behaviour and cognition. Proceedings of the 3rd International Workshop on Artificial Life and Artificial Intelligence, University of the Basque Country, December, 1994.* [TS]
(1995) On quantitative performance measures of robot behavior. *Robotics and Autonomous Systems* 15:107–34. [HJ]
Smolensky, P. (1986) Information processing in dynamical systems: Foundations of harmony theory. In: *Parallel distributed processing, vol. 1,* ed. D. E. Rumelhart & J. L. McClelland. MIT Press. [BS]
(1988) On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11:1–74. [RMF, aTvG, GO, JP, FvdV]
Steinhage, A. & Schöner, G. (1997) Self-calibration based on invariant view


https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press

recognition: Dynamic approach to navigation. *Robotics and Autonomous Systems* 20:133–56. [TS]
Sterelny, K. (1990) *The representational theory of mind.* Basil Blackwell. [GRM]
Stich, S. (1984) *From folk psychology to cognitive science.* MIT Press. [SRQ]
Swenson, R. & Turvey, M. T. (1991) Thermodynamics reasons for perception-action cycles. *Ecological Psychology* 3:317–48. [aTvG]
Tabor, W., Juliano, C. & Tanenhaus, M. (1996) A dynamical system for language processing. In: *Proceedings of the 18th Annual Meeting of the Cognitive Science Society.* Erlbaum. [aTvG]
Thelen, E. (1995) Time scale dynamics and the development of an embodied cognition. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [aTvG]
Thelen, E. & Smith, L. B. (1993) *A dynamics systems approach to the development of cognition and action.* MIT Press. [aTvG, MM]
Thom, R. (1983) *Mathematical models of morphogenesis.* Ellis Horwood. [aTvG]
Thurstone, L. L. (1928) Attitudes can be measured. *American Journal of Sociology* 33:529–54. [JRE]
Townsend, J. T. & Ashby, G. (1983) *The stochastic modeling of elementary psychological processes.* Cambridge University Press. [RAH]
Tsotsos, J. K. (1997) Intelligent control for perceptually attentive agents: The S* proposal. *Robotics and Autonomous Systems* 21:5–21. [HJ]
Tuller, B., Case, P., Mingzhou, D. & Kelso, S. J. A. (1994) The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance* 20:3–16. [arTvG]
Turing, A. (1936) On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2* 42:230–65. [aTvG]
(1950) Computing machinery and intelligence. *Mind* 59:433–60. [aTvG]
Turvey, M. T. (1990) Coordination. *American Psychologist* 45:938–58. [aTvG]
Turvey, M. T. & Carello, C. (1981) Cognition: The view from ecological realism. *Cognition* 10:313–21. [aTvG]
(1995) Some dynamical themes in perception and action. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [aTvG]
Vallacher, R. & Nowak, A., eds. (1993) *Dynamical systems in social psychology.* Academic Press. [aTvG]
Vallacher, R. R., Nowak, A. & Kaufman, J. (1994) Intrinsic dynamics of social judgment. *Journal of Personality and Social Psychology* 67:20–34. [JRE]
Van der Velde, F. (1995) Symbol-manipulation with neural networks: Production of a context-free language using a modifiable working memory. *Connection Science* 7:247–80. [FvdV]
(1997) On the use of computation in modelling behaviour. *Network: Computation in Neural Systems* 8:1–32. [FvdV]
Van Geert, P. (1994) *Dynamic systems of development: Change between complexity and chaos.* Harvester Wheatsheaf. [RAH]
(1995) Growth dynamics in development. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. MIT Press. [rTvG]
Van Gelder, T. J. (1990) Compositionality: A connectionist variation on a classical theme. *Cognitive Science* 14:355–84. [aTvG]
(1991) Connectionism and dynamical explanation. In: *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society.* Erlbaum. [aTvG]
(1993) What might cognition be, if not computation? *Cognitive Sciences Indiana University Research Report* 75. [CE]
(1995) What might cognition be, if not computation? *Journal of Philosophy* 91:345–81. [CE, aTvG]
(1997) Connectionism, dynamics, and the philosophy of mind. In: *Mindscapes: Philosophy, science, and the mind,* ed. M. Carrier & P. K. Machamer. Konstanz: Universitätsverlag Konstanz. [rTvG]
(1998) Some implications of connectionism for the philosophy of mind. In: *Philosophy and the Sciences of Mind. Proceedings of the Third Meeting of the Pittsburgh-Konstanz Colloquium in the Philosophy of Science.* [RMF]
(forthcoming) Bait and switch? Real time, ersatz time and dynamical models. In: *Proceedings of the Fourth Australasian Cognitive Science Conference, University of Newcastle, NSW Australia,* ed. R. Heath, A. Heathcote, C. Hooker & B. Hayes. University of Newcastle Press. [rTvG]
Van Gelder, T. J. & Port, R. (1994) Beyond symbolic: Towards a Kama-Sutra of compositionality. In: *Symbol processing and connectionist network models in artificial intelligence and cognitive modelling: Steps toward principled integration,* ed. V. Honavar & L. Uhr. Academic Press. [aTvG]
(1995) It's about time: An overview of the dynamical approach to cognition. In: *Mind as motion: Explorations in the dynamics of cognition,* ed. R. Port & T. van Gelder. Bradford Books/MIT Press. [CE, aTvG]
Von Bertalanffy, L. (1973) *General system theory: Foundations, development, applications.* Penguin. [aTvG]
Warren, W. H. (1995) Self-motion: Visual perception and visual control. In: *Handbook of perception and cognition, vol. 5: Perception of space and motion,* ed. W. Epstein & S. Rogers. Academic Press. [aTvG]
Wheeler, M. (1994) From activation to activity: Representation, computation and the dynamics of neural network control systems. *Artificial Intelligence and Simulation of Behaviour Quarterly* 87:36–42. [aTvG, MW]
Wildgen, W. (1982) *Catastrophe theoretic semantics: An elaboration and extension of René Thom's theory.* John Benjamins. [aTvG]
Wimsatt, W. C. (1972) Complexity and organization. In: *Proceedings of the Philosophy of Science Association, PSA 1972, vol. 2,* ed. K. Schaffner & R. S. Cohen. Reidel. [WB]
Zak, M. (1990) Creative dynamics approach to neural intelligence. *Biological Cybernetics* 64:15–23. [aTvG]
Zeeman, C. (1965) Topology of the brain. In: *Mathematics and computer science in biology and medicine.* Medical Research Council. [JP]



https://doi.org/10.1017/S0140525X9843173X Published online by Cambridge University Press