

# The Dynamical Challenge

ANDY CLARK

*Washington University*

Recent studies such as Thelen and Smith (1994), Kelso, (1995), Van Gelder, (1995), Beer, (1995), and others have presented a forceful case for a dynamical systems approach to understanding cognition and adaptive behavior. These studies call into question some foundational assumptions concerning the nature of cognitive scientific explanation and (in particular) the role of notions such as internal representation and computation. These are exciting and important challenges. But they must be handled with care. It is all too easy, in this debate, to lose sight of the explanatorily important issues and to talk at cross-purposes, courtesy of the (surprisingly) various ways in which different theorists often conceive the key terms. The primary goal of the present paper is thus a modest one: to begin to clarify just what is at issue and to highlight some of the most central and pressing concerns. In so doing, we may hope to develop a constructive framework for future debate. In addition, I try to open up a space of intermediate options—ways in which dynamical and representational/computational understandings may sometimes afford complementary (rather than competing) perspectives on adaptive success.

## 1. A NASCENT SCEPTICISM

These are exciting times for Cognitive Science. Once-unchallenged ideas concerning the nature of internal representational systems have been upset by the explosion of interest in connectionist and neural network models.<sup>1</sup> More recently still, even the bedrock notions of internal representation and computational explanation themselves have been subject to increasing critical scrutiny. In particular, several theorists concerned to do justice to the special nature of embodied intelligent systems have endorsed versions of a rather radical claim which goes something-like this:

The Radical Embodied Cognition Thesis: Structured, Symbolic, Representational and Computational views of cognition are mistaken. Embodied cognition is best studied using non-computational and non-representational ideas and explanatory schemes involving e.g. the tools of Dynamical Systems Theory.

Versions of this thesis can be found, for example, in recent work in developmental psychology (Thelen & Smith, 1994, Thelen, 1995), work on real-world robotics and autonomous agent theory (Smithers, 1994, Brooks, 1991), in philosophical treatments such as Wheeler (1994), and in some neuroscientific and neurobiological approaches such as Maturana & Varela (1980), Skarda and Freeman (1987). More circumspect treatments which nonetheless tend towards scepticism about computation and internal representation include Beer and Gallagher (1992), Beer (1995), Kelso (1995), Van Gelder (1995), Varela, Thompson and Rosch (1991), and essays in Port and Van Gelder (1995). Historical precedents for such scepticism are also in vogue and include especially Heidegger (1927:1965), Merleau-Ponty (1945:1962), and the work of J.J. Gibson and the ecological psychologists (e.g. Gibson, 1979).

The Radical Embodied Cognition Thesis constitutes, I believe, one of the most important and challenging developments in contemporary cognitive science. But it is a development whose genuine value is easily obscured by terminological misunderstandings (the word representation being an especially slippery case) and knee-jerk reactions (its just behaviorism, or, on the other side, cartesianism). The goal of the present paper is to clarify the nature of the genuine, open, empirical questions that are at issue and to develop a framework for constructive future debate. In addition, I shall try to open up a space of intermediate options—ways in which dynamical and representational tools may afford complementary (not competing) perspectives on adaptive success.

The strategy is as follows. I next (section 2) outline four ways in which the slippery term representation may be used. Of these four, only the last two constitute substantive, empirically significant options. Section 3 pursues some case studies of dynamical explanation and tries to identify a few guiding ideas. In section 4 I identify some assumptions that may seem to bridge the large *prima facie* gap between these ideas and the thesis of radical embodied cognition. I question these assumptions and, as a result, the relevance of the guiding ideas to the radical conclusions. What emerges, I hope, is a clearer sense of what really distinguishes the dynamical approach—namely a deep difference in explanatory emphasis (roughly, it is the difference between aiming to explain patterns and aiming to understand architectures). Section 5 pursues this difference, asks how the two projects interrelate, and suggests some ways in which they may ultimately prove complementary to one another. The concluding section draws these threads together to paint a picture of the main issues, and to (hopefully) clarify the space for future debate.

## 2. UNPACKING REPRESENTATION

The term internal representation has long stood firm as part of the basic infrastructure of cognitive scientific research and experimentation. Connectionists, it is true, diverged from tradition by putting their faith (for the most part) in a more implicit style of representing: they replaced the stable, simple, highly-manipulable symbols of classical Artificial Intelligence with numerical vectors and operations of vector completion and transformation. But though the computational profile differed, the basic commitment to a vision of intelligent behavior as involving the creation and use of internal representations remained inviolate.<sup>2</sup>

Contemporary critics of representational approaches tend, as we shall see, to distrust both classical and connectionist species of representationalism, though they typically believe the connectionists to be 'on the right track'—see e.g., (Wheeler, 1994), (Van Gelder, 1995), (Thelen & Smith, 1995) and others. Our immediate task, then, is to begin to clarify the (clearly quite general) sense of internal representation that looks to be at issue.

Let us begin with the weakest possible sense—one with which no one takes issue, except to note that it is so weak as to be totally uninformative. This is just the bare idea of internal state. It is agreed on all sides that flexible, adaptive, intelligent behavior often requires a creature to respond to current situations in ways informed by past experience, on-going goals and the like. Systems that merely react, in a pre-determined way, to immediate stimuli (that will always react the same way to the same stimulus) are unable to achieve this flexibility. What is needed is, at a minimum, the use of inner state to allow the agent to initiate and organize behavior without immediate environmental input, to anticipate future environmental inputs, and so on. In short, merely reactive agents are clearly inadequate to the full range of intelligent adaptive behaviors exhibited by biological organisms. Complex persisting and updatable inner state is thus at the heart of many (probably all) genuinely cognitive phenomena. This much must be common ground to both fans and sceptics about internal representation. The existence and importance of complex inner state is thus not at issue.<sup>3</sup>

Moving up a notch from the bare notion of inner state, we encounter the only slightly less vacuous notion of environmentally-correlated inner state. On its own, however, this requirement of correlation adds little to the bare idea of inner state. It would, after all, be almost miraculous if some kind of correlations between adaptively useful inner states and adaptively relevant environmental parameters did not exist. Moreover, correlation comes cheap (mappings can always be artificially defined) and is not necessarily even functionally illuminating. Thus Churchland & Sejnowski, (1992, pp. 185-186) describe a neural network in which certain hidden unit activities correlate rather nicely with the presence of edges—but the systemic role of these units is, in fact, not to do edge detection at all but to help extract curvature from shaded images. In sum, the bare idea of correlations between inner states and worldly features does not provide a substantial and illuminating sense for the term 'internal representation.'

The crucial moment in the transition to a genuinely substantive reading comes, instead, when we abandon the focus on mere inner state and/or correlation and replace it with a focus on the relation most usually glossed as 'standing-in.' All substantive notions of internal representation, I am willing to assert, have at their heart some idea of inner states (or processes) whose real functional role is to stand-in for other (usually extra-neural) objects, events, actions or states of affairs. Thus the philosopher John Haugeland insists (rightly) that representation-using systems are ones that achieve some kind of coordination with environmental features by the special method of having something else (in place of a signal directly received from the environment) stand-in and guide behavior in its stead (Haugeland, 1991, p. 144).

It is immediately apparent, however, that this notion of standing-in must now be treated with caution. For it can easily collapse back into the (too weak) notion of correlated inner

state. There are two ways to avoid this collapse, and they correspond to the only two substantive senses of internal representation that appear to be on offer.

The first (most popular, but still relatively weak) way to avoid the collapse is to add a proviso that standing-in requires not mere correlation but adaptively or functionally intended correlation. This usually amounts to requiring that the correlation be non-accidental and that the inner state is in some way consumed, within the larger system, for the specific information it therefore carries. For example, a population of neurons in the posterior parietal cortex of the rat may in this sense be said to act as stand-ins for head orientation, since other neural systems access their patterns of activity in ways that make functional (and evolutionary adaptive) sense if we assume they are accessing them so as to become apprised of the head-orientation information (in order, e.g., to run a radial maze—see McNaughton & Nadel, 1990, pp. 49-50). Let us call this the Weakest-Substantive-sense of internal representation, or weak internal representation for short.

This sense of weak internal representation is, I think, faithful to the bulk of contemporary neuroscientific usage—for such usage focuses on correlation but is willing (as in the Churchland & Sejnowski example) to reject some correlations as spurious or functionally impotent. There is, however, a final and still stronger sense of internal representation available. To bring it into focus, let us pause to rehearse a common (but initially puzzling) anti-representationalist response to the story so far.

The response goes something like this: all these readings of internal representation remain basically external—they show (at most) that it may sometimes help us to understand a system's inner organization if we identify certain inner states with information-bearing roles. But it does not follow, from this that those states are representations *for the organism itself*. This general kind of response can be found in e.g., (Brooks & Stein, 1993; Beer, 1995).

What can it mean? At first sight, it looks to be a version of the old and discredited idea that every representation needs an intelligent (maybe even conscious) user. But this, of course, is exactly the intuition that early work on automatic computing systems was supposed to defeat.<sup>4</sup> A more interesting and plausible reading, I suggest,<sup>5</sup> turns on the difference between inner systems that operate only so as to control immediate environmental interactions and ones that use inner resources to model the world so as to obviate the need for such continual environmental interaction.

An illustration will help. Consider two ways to solve the problem of finding your way out of a radial maze. One way relies (initially) on trial and error, but the system stores good solutions and (over time) supports successful maze-running. Activation of the maze-running routine is initiated by perceptual input that selects (by pattern association) a set of motor commands that will take the agent through the maze from a starting point thus specified—there can be many starting points, each having associated with it a learnt motor routine. In addition (let us assume) the agent can vary its responses according to its current goal states. If it is hungry, the sight of the maze will induce maze-running (there is food at the end!). If it is not hungry, the same perceptual input will have no effect.

This creature (call it Maze-runner 1) clearly exploits complex inner states whose functional role involves correlations with specific environmental features—hence it is a locus

of weak-substantive internal representation. But it is not really capable of modeling its world in the stronger sense mentioned above. To see what is missing, consider Maze-runner 2.

Maze-runner 2 is very like Maze-runner 1, but it includes some extra circuitry. This extra circuitry allows the agent to reason about maze-running off-line. Thus, confronted by a new way into the maze, the agent can deploy a tactic of vicarious exploration (see Campbell, 1974) to determine a viable route in advance of actual physical action. To support such functionality, the system uses distinct inner states as stand-ins for distinct features of the maze, and is set up (by learning, evolution or hand-coding) so that the relations between these inner states mirror the actual relations (of distance, accessibility, etc.) between real-world maze-features. Notice, then, that all that need ultimately distinguish such a case from any similarly articulated case of weak-substantive internal representation is the capacity to access such inner structures off-line and thus to support planning and problem-solving in the absence of rich on-going environmental exchange.

This description of strong internal representation is deliberately vague concerning the actual mechanisms involved (for a worked example involving the capacity to strongly model potential motor activity, see Grush, 1995). For it does not matter what the mechanisms are (neural nets, object-oriented programs, expert systems, etc.) as long as they display certain key features. First, they must involve inner states or processes whose functional role is to coordinate the system's activity with its world (no mere correlations). Second, we must be able to identify specific inner states or processes with specific representational roles—we must be able to isolate, within the system, the inner parameters or processes that stand-in for particular extra-neural states of affairs (otherwise we confront only complex inner state implicated in successful agent-environment coordination). And lastly, the system must be capable of using these inner states or processes so as to solve problems off-line, to engage in vicarious explorations of a domain, and so on. It is this last capacity that distinguishes the genuine model-using agents from the rest.<sup>6</sup> Strong internal representation is thus of a piece with the capacity to use inner models instead of real-world action and search. Inner states and processes that function as stand-ins in such models are, I suggest, genuinely representations for the agent and not simply useful glosses imposed from outside.

### 3. THE DYNAMICAL CHALLENGE

The Radical Embodied Cognition Thesis (see section 1) is motivated, in part, by a number of recent demonstrations spanning a variety of disciplines and approaches including developmental psychology, robotics and autonomous agent theory, and the general study of dynamic pattern formation. The common ground of these various investigations is (very roughly) the idea that certain target phenomena—including some cognitive and psychological ones—are best understood as the emergent products of the complex, often non-linear and temporally rich, interplay between a variety of forces. This interplay can be wholly internal or (more frequently) can involve, as 'equal partners,' internal, bodily, and environmental factors. Where target phenomena depend on such complex interactions (internal or

otherwise), they are best explained, it is argued, using alternative, non-representational tools (such as those of pure dynamical systems theory) -tools suited to the study of emergent phenomena in de-centralized and self-organizing systems.

To get the flavor of these ideas in action, let us briefly (too briefly, but see Clark, 1997, for a fuller treatment) examine three representative cases, beginning with a basic physical example.

### Case One: The BZ Reaction

In the Belousov-Zhabotinsky (BZ) reaction, four reactants (malonic acid, potassium bromate, cerium sulfate and sulfuric acid) are pumped into an agitated tank. After a time, the flow of reactants is steadily increased. All the while, the concentration of one of the products of the reaction (the bromide ion) is measured as a function of time. The result is a fascinating pattern of unfolding in which the oscillation periods of the target ions first increase by doubling and redoubling the original periods and then (at a certain rate of flow of reactants) step into the complex but deterministic patterning known as chaos (see e.g., Gleick, 1987). The discovery of this complex, time-critical patterning is significant as a demonstration of one characteristic feature of much dynamical modeling: the use of low dimensional but temporally rich descriptions to display key organizational properties of highly complex systems. The full dynamics of the system of four reactants and their many (20+) products is immensely complex. The beauty of the dynamical description is that it reproduces the trajectory of the overall, complex system in its high-dimensional space by using the time-linked values of the bromide ion concentration as an effective but low-dimensional image of the systems unfolding. The idea is thus that even though a single degree of freedom does not give a full picture of the actual complicated dynamics, the time evolution of this single variable is affected by the other degrees of freedom and thus contains information about the whole dynamical process (Rueger & Sharp, 1996, p. 104).<sup>7</sup>

The trick of a good dynamical systems analysis in such cases, is to find and plot the low-dimensional description that captures patterns of change over time of special interest to an experimenter. This is the first of three guiding ideas I want to highlight. The upshot is what Rueger & Sharp call simple theories of messy systems—illuminating accounts of behavior so complex that it might otherwise appear totally random. Non-linear dynamics (in which a smooth change applied to one value can yield a large jump or shift, at a critical point, in a system's overall behavior) and chaotic unfoldings constitute ideal targets for such forms of understanding.

More generally, dynamical systems theory comprises a large set of mathematical and topographic tools that together provide a powerful vocabulary for the description and analysis of patterns of change in system parameters over time. These tools include mathematical techniques for the construction, comparison and analysis of time-sequence data (such as the use of delay coordinates—see Simonyi, 1982) and more qualitative constructs that help capture the basic topological features of the state spaces defined by the chosen parameters. The core ideas are thus the notion of a state space, the idea of a set of possible trajectories (the flow) through that space and the use of continuous or discrete mathematics to describe the laws that fix the shapes of the possible trajectories. Certain regions of a state

space may exhibit notable properties. An attractor is a point or region such that any trajectory passing close by will be drawn into the region (the area of such influence being known as the basin of attraction). A repeller is a point or region that deflects incoming trajectories. A bifurcation is a point at which a small change in parameter values can re-shape the flow within the state space and yield a new landscape of attractors, repellers and so on. Dynamical systems approaches thus provide a set of mathematical and conceptual tools that help display the temporal and spatial order in the evolution of specific systemic parameters. Our next example illustrates this and introduces a second guiding idea viz. the use of collective variables.

### Case Two: Rhythmic Finger Motion

In the case of the BZ reaction, a useful low-dimensional description was achieved by focusing on one actual product of the ongoing reaction—the concentration of bromide ions. Sometimes, however, the search for powerful low-dimensional descriptions requires the experimenter to actively define new collective variables. These are variables that do not track properties of simple physical parts but instead track higher-level properties that may involve e.g., relations between measured values of physical parts. Thus consider the case (Kelso et al, 1981), (Kelso, 1995, Ch.2) of rhythmic finger motion.

Human subjects, asked to move their two index fingers at the same frequency in a side-to-side wiggling motion, display two stable strategies. Either the fingers move in phase (the equivalent muscles of each hand contract at the same moment), or exactly anti-phase (one contracts as the other expands). The anti-phase solution, however, is unstable at high frequencies of oscillation—at a critical frequency it collapses into the phased solution.

How should we explain and understand this patter of results? One strategy (a version of the BZ strategy displayed above) is to seek a more illuminating description of the behavioral events. To this end, Kelso and his colleagues plotted the phase relationship between the two fingers. This variable is constant for a wide range of oscillation frequencies but is subject to a dramatic shift at a critical value—the moment of the anti-phase/phase shift. Plotting the unfolding of the relative phase variable is plotting the values of a collective variable since relative phase is determined by a relation between the behaviors of more basic system components (finger motions). The values of this collective variable were observed to be fixed by frequency of motion, which thus acts as a so-called control parameter. The dynamical analysis is then fleshed out by the provision of a detailed mathematical description—a set of equations displaying the space of possible temporal evolutions of relative phase as governed by the control parameter. This description fixes, in detail, the state space of the system: the attractors, repellers, bifurcation points and so on. Haken et al (1985) uses such an analysis to display the different patterns of coordination corresponding to different values of the control parameter. Some noteworthy features of the resulting model are 1) its ability to account for the observed phase transitions without positing any special switching mechanism—instead, the switching emerges as a natural product of the normal, self-organizing evolution of the system, 2) its ability to predict and explain the results of selective interference with the system (as when one finger is temporarily forced out of its stable phase relation), and 3) its ability to generate accurate predictions of e.g.,

the time taken to switch from anti-phase to phase. (For a nice review of the model, see Kelso, 1995, p. 54-61).

The dynamical explanation is thus perched midway between what, to a more traditional cognitive scientist, may at first look like a ('mere') description of a pattern of events and a real explanation of why the events unfold as they do. It is not a mere description since the parameters need to be carefully chosen and the resulting model has predictive force: it tells us enough about the system to know how it would behave in various non-actual circumstances. But it differs from more traditional cognitive scientific explanations in that it uses collective variables to abstract away from the behavior of individual systemic components. This tendency towards collective-variable style abstraction constitutes the second guiding idea I wish to highlight.

### Case Three: Treadmill Stepping

Consider the phenomena of learning to walk. Thelen and Smith (1994) show, quite convincingly, that these phenomena (in human infants) demand explanations which invoke a multiplicity of factors spanning brain, body and local environment. Such explanations differ markedly from certain traditional schemes in which such progressive changes are depicted as the inexorable playing out of a set of prior instructions encoded in e.g. a genetically specified central pattern generator or neural control system (Thelen & Smith, 1994, pp. 8-20, 263-266). The difference lies principally in the way the problems themselves are conceived, and the resulting multidimensional nature of the solutions viz. "a multi-dimensionality in which the organic components and the context are equally causal and privileged" (Thelen and Smith, *op. cit.*, p. 17).

In the case of learning to walk, Thelen and Smith address the following pattern of developmental transitions:

1. Newborn infants, held upright, produce efficient, coordinated stepping movements.
2. These movements disappear at about 2 months of age.
3. The movements reappear at about 8 to 10 months of age when the infants begin to support their weight on their feet.
4. Finally, at about 12 months, the first independent steps are observed.

One explanation of these regular transitions would be to posit a detailed developmental plan or blueprint encoded in the central nervous system which (perhaps for reasons steeped in the idiosyncratic evolutionary history of the species) constrains the infant to display this particular sequence of intermediate forms. By contrast, Thelen and Smith argue, convincingly, that "walking...is not controlled by an abstraction, but in a continual dialogue with the periphery" (*op. cit.*, p. 9). To illustrate this, the authors report some striking data which shows how stepping motions are "soft assembled" out of a complex combination of neural, bodily and environmental influences.

For example, Thelen and Smith report that stepping motions can be induced even between stages 2 and 3 above (i.e. during the period in which held erect stepping is absent) if the infant is held upright in warm water. Moreover, non-stepping 7-month infants held upright on a slow motorized treadmill performed highly co-ordinated alternating steps,

adjusted step speed to compensate for increased treadmill speed, and even made asymmetrical leg adjustments to maintain rhythmic stepping on two belts driving opposing legs at different speeds! Other environmental manipulations (such as adding weights to the legs) were able to inhibit stepping in infants who normally displayed it. Such results (described in detail in Thelen and Smith, *op. cit.*, Ch. 1 & 4) show that the behavioral repertoire of the infants is highly sensitive to bodily and environmental parameters such as the effective weight of the legs. Such observations lead the authors to conclude that "there is no essence of locomotion either in the motor cortex or the spinal cord. Indeed, it would be equally credible to assign the essence of walking to the treadmill than to a neural structure..."(Thelen and Smith, *op. cit.*, p. 17).

The treadmill stepping task, thus provides an especially useful window onto the dynamical construction of infant walking, as it highlights the complex and subtle interplay between intrinsic dynamics, organic change and external task environment. In fact, the treadmill looks to be acting as a real-time control parameter that prompts the phase shift, in the 7 month olds, from non-stepping to smooth alternating stepping motions. To test and refine this hypothesis, Thelen & Ulrich, (1991) focused on the alternation of steps (the treadmill response that seemed most suggestive of mature locomotion patterns). They used the relative phasing of the two legs as a collective variable, since this distinguishes true alternation from other possible leg actions that have less in common with mature locomotion strategies (in mature stepping the two legs are at .5 relative phase i.e., 180° out of phase—one leg is 180° through its cycle when the other initiates its motion). In their study, Thelen & Ulrich plotted relative phasing in treadmill stepping in infants from one to eight months of age. The authors plotted month-to-month performance (infant by infant) and also observed the effects of varying treadmill speed both between and during individual trials. When held on the treadmill, infants could display a variety of responses (no action, single step, double step, parallel step, alternating step). In the early months, all these were frequently observed. But alternation steadily became the preferred response: in dynamic terms, the multi-stable states of single, double, parallel and alternating were replaced by the singular state of alternation, the attractors both dissolving and evolving over the first 8 or 9 months (Thelen & Smith, 1994, p. 103). With a large body of data to hand, the next step was to try to discover possible control parameters (organic or environmental) to account for the detailed profile of the onset of treadmill stepping. Out of a host of options, one factor coordinated especially well with the data: the orientation of leg and foot in relation to the treadmill. Poor stepping was correlated with high degrees of leg flex and inward rotation of the foot. Good stepping was correlated with flat-foot belt contact (rather than toe contact). And flat-foot contact itself was negatively correlated with high leg-flex. Thus imagine that the leg, when stretched out, acts like a spring—at full stretch, the energy imparted to the spring is released and swings the leg forward. Receptors in the muscles respond to the uncoiling of the leg and this information is used to control relative phasing. On such a scenario (see Thelen & Smith, *op. cit.*, p. 111-12) it is crucial that the leg be fully stretched back, by the treadmill action, to initiate alternating stepping. Still treating infant legs like mass springs, we can now see that if the intrinsic tension of the leg is too high, the treadmill won't manage to produce the back-stretch needed to activate the receptors and

initiate phased stepping. Increasing treadmill speed will, in some borderline cases, yield the necessary stretch. Young infants, it is observed, have high flexor bias in the legs—they are coiled up and only relax over several months. The authors conclude that:

...the relative flexor (very tight) or extensor (more loose) tendencies of the legs, in this case as indexed by several postural characteristics, acted as the control parameter to engender the shift into stable alternate stepping. As a control parameter, flexor tone constrained the interacting elements, but did not prescribe the outcome in a privileged way...the emergence of coordinated treadmill stepping must be a multi-determined process. While it seems likely that the pathways essential for treadmill-stepping pattern production can function by 1 month of age, central neural pattern generation is likely not the developmental control parameter in this case. Rather, the behavior itself emerges only when the central elements cooperate with the effectors—the muscles, joints, tendons—in the appropriate physical context (Thelen & Smith, 1994, p. 113).

The third (and final) guiding idea to emerge from our case studies is thus the image of 'soft assembly' in an extended (brain/body/world) system— an image that leads us to depict development as the successive creation and dissolution of attractors in a distributed system whose organic and environmental components are changing over time.

#### 4. REPRESENTATION REVISITED

How does all this bear on the issues concerning internal representation and computationalist theorizing? The question is difficult because, at first sight, the relevance can appear somewhat marginal. The specific problems addressed seem far removed from more traditionally cognitive topics such as planning, speech, story-understanding and so on. The vocabulary and methodology do indeed seem very different. But it is hard to avoid the suspicion that these differences stem largely from this difference in topic (basic motor skills versus 'real thinking'). It is thus not immediately clear how the ideas bear on the much more general thesis of radical embodied cognition. Nonetheless, there can be no real doubt concerning the authors' intentions. We read, for example, that:

**Explanations in terms of structure in the head**—beliefs, rules, concepts and schemata—**are not acceptable**...Our theory has new concepts at the center—nonlinearity, reentrance, coupling heterochronicity, attractors, momentum, state spaces, intrinsic dynamics, forces. These concepts are not reducible to the old ones. (Thelen & Smith, 1994, p. 339). (My emphasis).

We posit that development happens because of the time-locked pattern of activity across heterogeneous components. We are not building representations of the world by connecting temporally contingent ideas. **We are not building representations at all! Mind is activity in time**...the real time of real physical causes (Thelen & Smith, 1994, p. 338). (My emphasis).

The thesis here is that the human brain is fundamentally a pattern-forming, self-organized system governed by non-linear dynamical laws. **Rather than compute**, our brain dwells (at least for short times) in metastable states... (Kelso, 1995, p. 26). (Second emphasis mine).

Linking the investigations and guiding ideas just rehearsed and the more general and radical thesis are, I suggest, three bridging assumptions. First, there is an (often explicit) hypothesis of continuity<sup>8</sup>—a claim to the effect that all of cognition is continuous with its motor and developmental foundations, and hence that the shape of solutions (emergent, situated, soft-assembled) to these problems will be recapitulated in all supposedly higher cognition domains (see e.g., Thelen & Smith, 1994, p. xxiii). Second, there is an assumption about the ‘objectivist’ nature of any putative internal representations. And third, there is an assumption about the kind of behavior control systems implicated in computational accounts. All three bridging assumptions bear discussing, and I will address each in turn.

First, the hypothesis of continuity. The problem here (a familiar one to evolutionary biologists) is that the simple notion of continuity covers a remarkably wide variety of possible linkages, pathways and commonalities. In the genetic case these include the ideas of fairly smooth evolution with constant function (as in the case of the eye), smooth structural evolution with radical functional change (as in the case of the role of thermo-insulating feathers in enabling flight) and symbiosis, in which various parts evolve quite separately and are put together at a later date so as to usher in some brand new functionality. For detailed discussion of all these cases see Ridley (1985) p. 35-41. In a similar manner the ontogenetic continuity posited by the dynamicist surely exists in some form or other. But it hardly follows from this that the guiding ideas and vocabulary suited to e.g., the explanation of early motor development will apply across the board.

Consider, for example, the strongest representationalist thesis outlined in section 2 above. According to this thesis, we sometimes solve problems by exploiting inner states that are designed (by learning or evolution) to function as off-line stand-ins for features of our real-world environment. In such cases (Maze-runner 2 was our example), we temporarily abandon the strategy of directly interacting with our world so as to engage in more vicarious forms of exploration. It is certainly possible that such off-line problem-solving is perfectly continuous in a sense with various on-line highly environmentally interactive, motor control strategies. Thus Grush (1995) describes a piece of circuitry whose principal role is the fine-tuning of on-line reaching. The circuitry, however, involves the construction of an emulator loop that predicts sensory feedback in advance of the signals arriving from the bodily peripheries. This loop, once in place, can later support the additional functionality of fully off-line deployment, allowing the system to rehearse motor actions entirely in its imagination. Such a case shows both a profound continuity between smooth motor control strategies and higher cognitive capacities such as off-line reasoning and imagination, and (simultaneously) a profound *discontinuity* in that the system, as described by Grush, is now using specific and identifiable inner states as full-blooded stand-ins for specific extra-neural (in this case bodily) states of affairs. These are internal representations in the very strong sense defined in section 2. At such times the system is not continuously assembling its behavior by balancing ongoing neural, bodily and environmental influences. It is, instead, modelling and representing its world. We thus preserve a kind of architectural continuity, while abandoning the guiding idea of soft assembly in an extended system (for a more detailed treatment of the implications of this case, see Clark & Grush, to appear).

The second bridging assumption concerned the nature (content) of any putative internal representations. Here, it looks as if the target of a great deal of dynamicist scepticism is not internal representation *per se* so much as a particular type of internal representation *viz* what are sometimes called objectivist representations—the kind that might feature in a detailed, viewpoint-independent, map-like model of some aspect of the world. Notice, then, a second (and I believe, highly significant—see Clark, 1995, Clark, 1997) way in which higher level cognition may be innocently continuous with its motor and developmental roots. It may be continuous insofar as it involves internal representations (weak or strong) whose contents (unlike detailed objectivist representations) are heavily geared towards the support of typical or important kinds of real-world, real-time action. Such contents may (as in the previous example) sometimes be manipulated off-line—but they are nonetheless types of content (what I elsewhere call action-oriented contents) that are especially suited to the control and coordination of real action in real time. Cognition, on this model, need not always be actually interactive (involving brain, body, and world as equal partners). But the inner economy has still been sculpted and shaped by the real-time, task-specific, interactive needs of the organism.

Much dynamicist scepticism, on closer examination, looks to address only the specific notion of objectivist (detached, action-independent, highly-detailed, static, general-purpose) internal representation. Thus Thelen & Smith (1994, p. 37-44) question all those ideas, suggesting instead that we treat knowledge as an action-guiding process continually organized against a contextual backdrop that brings forth its form, content and use. The same set of emphases characterize Varela's notion of enaction in which cognitive structures are said to emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided (Varela, Thompson & Rosch, 1991, p. 173). To mark the difference, Varela, Thompson and Rosch define a sense for the term 'strong representation' that specifically associates strong representation with the disputed idea of an inner recapitulation of the objective features of a 'pre-given world' (op cit p. 148). Such a sense of strong representation, it should be clear, places the emphasis on the kinds of content involved rather than the functional role of the inner states. In a similar vein, Agre (1995) notes the importance of what he calls indexical-functional representations (such as 'a few feet straight ahead')—these are ideal for the cheap control of individual action and are to be contrasted with objectivist map-like representations such as "at latitude 41, longitude 13". The point I want to stress is just that many disputes in this area thus look to concern the content, not the existence, of inner states whose role is to stand-in (in either the weak-substantive or strong sense) for adaptively important extra-neural states of affairs. At a minimum, the mere fact (if it is a fact) that biological agents rely heavily on indexical-functional, action-oriented and context-responsive kinds of knowledge does not, in itself, undermine the idea that such knowledge may itself be internally encoded in ways that allow us to gain illuminating insight into the functional organization of the agent by seeking a representationalist understanding— that is to say, by associating quite specific inner states and processes with the tasks of maintaining and manipulating such information. Of course, there is no guarantee that this will be the case—the knowledge may be so widely and complexly distributed between various inner subsystems and bodily and environmental factors that even the

weakest substantial representational analysis fails. The point is just that the debate over internal representation really turns on just these further issues (concerning complexity and distribution—see section 5 following) rather than on issues concerning the kinds of content encoded by any posited representations. Arguments against specifically objectivist representational contents should thus play no role in the more general discussion between the friends and foes of internal representation.

Lastly, there is a bridging assumption concerning the type of inner control strategies implicated in broadly representationalist/computationalist accounts. The assumption, roughly, is that computational models involve the storage and use of complex inner control structures that plot, in explicit detail, all the values and settings of all the physical parameters involved in a given action. Something like this assumption would help explain why Thelen & Smith repeatedly associate the idea that the brain is a computational device with ideas about detailed advance blueprints for behavior, complete with internal clocks, specifications of all relevant parameter settings (joint-angle coordinates, muscle flexing patterns, etc.) for the limbs, and so on (Thelen & Smith, 1994, p. 75—see also *op. cit.*, p. xix, 62-63, 264, 71). They then contrast this vision of highly detailed, complete neural instruction sets with ideas of collective states, phase shifts and control parameters, as discussed in the previous section. Certain preferred collective states of the system are depicted as synergetic wholes that can be brought forth (but not programmed) by the action of some control parameter (such as frequency of motion in the rhythmic finger motion case, and flexor tone in the treadmill stepping case). A typical dynamical story thus focuses not on neural instruction sets but on some stable collective variable compressing the possible combinations of the structural components in an energetic and task content (*op. cit.*, p. 63)—a context that can build in not only inner facts, but local environmental structure and so on. Kelso's description of the brain itself as "not a computing device but a pattern-forming, self-organized system" (Kelso, 1995, p. 26) has the same flavor. The contrast is between systems whose behavior is fixed by complex encoded instruction sets and ones whose behavior emerges as a sequence of temporarily stable states (tracked by collective variables) in a complex system with richly inter-dependent intrinsic dynamics. The slogan may be 'patterns without programs.' But the target of the criticism is the more restricted idea that we use complex neural instruction sets to force orderly behavior from multiple muscles, links, joints, etc. Such detailed forcing is not necessary, it is claimed, because the system self-organizes into a smaller set of preferred states whose flux may be controlled by the action of some simple parameter. (It is a little as if the computationalist, faced with the problem of moving a crowd from A to B, were to encode an instruction for each persons trajectory, whereas the dynamicist simply finds a control parameter (maybe increasing the heat on one side of the crowd) that then exploits the intrinsic responses of those closest to it, whose motion in turn entrains the movements of their near neighbors, until the crowd moves as a unified whole in the desired direction).

This is an important and fascinating shift in emphasis, to be sure. But does it really amount to a rejection of the idea that the brain is a computer? I suggest two reasons to be doubtful. First, there is no necessary commitment on the part of the computationalist to the idea of highly detailed or complete instruction sets (indeed, Thelen & Smith recognize this

point in their comments on schemata: general movement plans that do not dictate specific kinetic details—*op. cit.*, p. 75). Thus consider the very idea of a program for performing some task. The basic idea is indeed that of a set of instructions that, when followed, lead to a solution. But what is the difference between a set of instructions (a recipe) and a mere force which, if applied, brings about a result? The heat applied to a pan of oil, at a critical value, leads to the emergence of convection rolls (see Kelso, 1995, pp. 6-8). Yet, it is not, intuitively, a recipe for such currents. One difference is that a recipe is written in some language of arbitrary symbols. But another is that the guiding parameter, in the case of the heated oil, seems too simple and unarticulated to count as a program. It is more like plugging a computer in than running a piece of software, as one of my students usefully remarked. Nonetheless (and here is where things get murky), it seems clear that genuine programs can vary markedly in complexity. A short piece of software, written in a high-level language, will not itself specify how or when to achieve many sub-goals—these tasks are ceded to built-in features of the operating system or to the activity of a cascade of lower-level code. Moreover, a program can perfectly well assume some necessary backdrop of environmental or bodily structures and dynamics. Jordan et al (1994) describes a program for the control of arm motions, but it is one that assumes (for its success) a lot of extrinsic dynamics such as the mass of the arm, the spring of muscle and the force of gravity. My claim, then, is that we here confront not a dichotomy (programmed versus unprogrammed) but a continuum—the less detailed the specification required (the more work is being done by the wider intrinsic physical dynamics of the system), the less value there is in treating the neural contributions as any kind of a program.

Now it may be, of course, that so very much is done by the synergetic dynamics of the body-environment system that the neural contributions are indeed best treated, at all stages, as the application of simple forces to a complex but highly inter-animated system whose intrinsic dynamics then carry most of the load. But less radically, it may be that basic motor activity simply requires less in the way of detailed inner instruction sets than we might have supposed, courtesy of the existence of a small set of preferred collective states such that successful behavior requires only e.g., the setting of a few central parameters such as initial stiffness in a spring-like muscle system and so on. Such sparse specifications may support complex global effects without directly specifying joint-angle configurations and the like.

The lack of a particularly detailed kind of neural instruction set does not then, establish the total absence of stored programs. Such a characterization is compelling only at the most extreme (and perhaps basic motor-control specific) end of a genuine continuum. Between the two extremes lies the interesting space of what I elsewhere (Clark, 1997) call 'partial programs'—minimal instruction sets that maximally exploit the inherent (bodily and environmental) dynamics of the controlled system. The real moral of much actual dynamical-systems-oriented research is, I suspect, that it is in this space that we may expect to encounter many of nature's own programs.

There is, of course, a second way in which a program differs from an applied force. A program is, in some surprisingly elusive sense, constituted by a set of commands or descriptions couched in some kind of code. This notion of a code is intuitively one that involves

ideas of arbitrary symbols and some capacity for the concatenation and recombination of such symbols to express new contents. It is unclear, however, whether we should assimilate the general vision of the brain as a computational device to this narrower vision of the brain as a locus of such compositional coding schemes. Many connectionists who do not subscribe to the use of traditional compositional coding schemes typically regard themselves as nonetheless investigating the space of potential computational strategies that might figure in biological cognition (see e.g., McClelland, Rumelhart et al, 1986, Churchland, 1989, Smolensky, 1988, Elman, 1994). One reason for this is clear. There is a sense of ‘computation’ that is tied not so much to the stored program image as to the idea of systems that effect automatic, semantically-sensible transformations between internal representations<sup>9</sup>. Recall now the weak-substantive sense of internal representation outlined in section 2. In this sense, we find internal representations whenever we can identify an isolatable inner state or process with the functional role of standing-in for specific, usually extra-neural, states of affairs. Such states may, however, be found in connectionist systems that lack many of the usual features of classical code and symbol devices (for a review, see Clark, 1993, Ch.3). In such cases the failure of the stored, coded, compositional program idea need not undermine the vision of the system as broadly computational in the sense just described. If it is indeed the case (and this is open to question- see e.g., essays in Harnad, 1994 for some discussion) that we find computation whenever we find automatic, semantically-sensible transitions between internal representations, then the question of whether the brain really computes reduces to the prior question about internal representation. (My own view, though I will not argue for it here, is that the notion does thus reduce—hence the emphasis, in the present treatment, on issues concerning internal representation).

All three bridging assumptions thus face serious difficulties. As a result, I can find no clean and compelling route from the various guiding dynamical ideas to the strong conclusions expressed in the Radical Embodied Cognition Thesis. Instead, what looks to be strongly supported is a pair of very reasonable and important strictures that may usefully inform both computationalist/representationalist and ‘pure dynamical’ investigations. They are:

1. Beware of putting too much into the head. Adaptive behavior emerges from a complex balancing act that incorporates neural, bodily, and environmental influences.
2. Beware of narrow visions of the form and content of putative internal representational systems. Such systems may involve indexical-functional (action-oriented) contents and may not require expression in the form of compositional codes and classical programs.

This more conservative message, however, should be coupled with the realization that dynamical approaches are genuinely successful in forcing a much-needed re-examination of the explanatory goals, methods and expectations that inform much cognitive scientific research. What our discussion indicates, I suggest, is that the real challenge lies not in the supposed implications for notions of representation and computation but in the ideas concerning the dense spatial and temporal interplay between neural, bodily and environmental factors, and the kind of tools and construct (collective variables, control parameters and so

on) needed to distill illuminating order from this mass of burgeoning complexity. The true dynamical challenge thus lies in the imperative to shift our primary focus from the componential detail of internal architectures to the fine-grained structure of bodily and environmentally extended patterns. It is to this challenge that we now turn.

## 5. PATTERN, ARCHITECTURE & EXPLANATION

The notions of patterns and of self-organization dominate much dynamical research. We have seen these notions in action in all the case studies reported in section 3. Closely related notions include those of coupling and of circular causation. These latter notions are especially useful in describing and analyzing the patterns that emerge in ongoing organism-environment interactions. Two sources of variance (the organism and the environment) may be depicted as a single coupled system whose evolution is specified by a set of interlinking equations. Thus consider two pendulums mounted close together on a wall. The pendulums tend to become swing synchronized over time courtesy of vibrations running through the wall (see Salzman, 1995). This synchronization admits of an elegant dynamical explanation in which the motion equation for each pendulum includes a term that is fixed by the other's current state. As each system (each pendulum) moves through its state space, it effectively alters the shape of the state space of the other system. These altered dynamics simultaneously transform the state space of the first system. The larger coupled system thus displays a kind of circular causation (see Merleau-Ponty, 1942, Varela et al, 1991, Kelso, 1995, Ashby, 1956) in which each subsystem is continuously influencing, and being influenced by the other. This kind of complex interplay is what I call continuous reciprocal causation (Clark, 1997)—it can lead, in many cases, to large-scale emergent behaviors whose quality and complexity far exceeds that which either subsystem could display in isolation.

The notion of coupling (see e.g., Beer, 1995, Van Gelder, 1995) thus provides a mathematically elegant way to display and understand the sometimes very complex interplay that characterizes systems that both affect and are continuously affected by their surroundings. Biological systems, it seems clear, are a case in point. Perception and action, as Merleau-Ponty long ago pointed out (see also Varela et al, 1991 and, in a computational vein, Ballard, 1991) are bound together in just such an intimate loop. Our perceptions guide actions that alter perceptions that guide further actions and so on. Moreover, this same kind of circular complexity may characterize even purely internal relations such as those that obtain between distinct neural and bodily subsystems—see e.g., Cohen's (1992) work on heterarchical control structures linking brain, spinal cord, muscles and limbs, or Knierim and Van Essen's (1992) work on hierarchical influences in visual processing. Such dense and reciprocal interactions pose special analytic problems—ones that may well require the use of dynamical tools to display salient low-dimensional regularities in the complex web of interactive influence.

When such complex interactivity is present, it seems likely that some kind of explanatory priority should be given to attempts to discover (via the use of collective variables, coupled dynamical equations, control parameters and so on) the (perhaps temporarily) sta-

ble global states that such complex exchanges support. Notice, however, that complexities that turn *only* on ongoing organism-environment interactions cannot support a case against strong representation, since this (by our definition) crucially involves a potential de-coupling between the inner representational system and the stream of immediate environmental input. The immediate point, however, is that there is no inherent opposition between the project of giving explanatory priority to the understanding of patterns in the behavior of coupled organism-environment systems and the project of understanding how specific inner architectures contribute to adaptive success. It is agreed on all sides that certain dynamical stories, qua low-dimensional takes on a higher dimensional reality, may obscure (or at least fail to reveal) important details concerning the actual physical mechanisms that are at work. This is just the price we pay for the explanatory insights they provide concerning the higher level patterns that emerge in complex, self-organizing systems. Thus two prominent theorists studying the BZ reaction using dynamical tools comment that such approaches “aim at revealing the mathematical essence of the experimental periodicity” and are “of limited use for understanding the chemical mechanism that generates the complexity” (Gyorgyi & Field, 1993, p. 55). Similarly, Thelen & Smith, immediately after their illuminating discussion of collective variables and control parameters in treadmill stepping, note that the resulting story is “clear about the general processes of change through the loss of stability of coherent dynamic organization, but completely uninformed about the more precise mechanisms of changing attractor stability “ (Thelen & Smith, 1994, p. 129). In response to this need, Thelen & Smith go on to pursue issues concerning the dynamics of the underlying neural organizations. Kelso, likewise, insists on a tripartite scheme in which full understanding requires an analysis of the task itself, a high-level dynamical account pitched at the collective variable level, and an account pitched at the component level—see Kelso (1995, p. 66). He also cautions, importantly, that what counts as a component or a collective variable depends to some degree on our current explanatory interest. Non-linear oscillators, he notes, may be treated as components for some purposes. Yet such oscillation is itself a collective effect that emerges from the interactions of still simpler components. Randall Beer, in a series of careful and progressive attempts to understand the operation of neural network controllers for embodied action in insect-like agents has also stressed the need to pursue the dynamical understanding all the way down. His project targets the detailed dynamics of individual neuron-like units, then coupled pairs of units, then pairs of units coupled with simple bodies, and so on (see Beer, 1995). There is, it seems, a quite general recognition that the explanatory aspirations of cognitive science require us to move, at some point, beyond the collective-variable style depiction of gross emergent patterns and into the realm of inner architectures, components and organizations. There is thus a clear space for complementary investigative activity spanning high-level dynamics and more detailed, component-oriented research. The question is really one of explanatory priority, with fans of embodied and situated approaches advocating greater attention to the large-scale, ecologically crucial patterns to which inner neural organizations contribute as more or less equal partners with bodily and environmental factors.

When the more inner oriented research is conducted using dynamical constructs and tools, however, an interesting possibility emerges. Even if there are no simple, static fairly

local states of the brain that can usefully be associated with representational roles (due, perhaps, to the presence of dense reciprocal connectivity and circular causal influences), there might still be higher level dynamical regularities that are playing just such an adaptive role. The opportunity thus exists to use the distinctive analytic tools of dynamical theorizing (collective variables, attractors, trajectories and so on) to distill potential representational vehicles from the burgeoning complexities of internal neural activity.

Such, indeed, is already the case in active research in artificial neural networks. Such networks afford multiple potential vehicles for internal representations, including states of distributed activity, attractors defined within the state space of units and weights, and trajectories in the state space (see e.g., Smolensky, 1988, Elman, 1991). Other dynamical entities such as chaotic attractors, limit cycles, and bifurcation structures are equally capable of playing the adaptive roles characteristic of either weak or strong representation, as described in section 2. Moreover, as dynamicists increasingly turn their attention to more traditionally cognitive and 'representation-hungry' (the term is from Clark & Toribio, 1994) domains, such as long-term planning and decision-making, the need to press such vehicles into representational service grows stronger—this is clearly seen in e.g., the various essays in Port & Van Gelder (1995), Nehmzow & Smithers (1991), Amit (1989), Van Gelder & Port (1994), Petitot (1985), Miller & Freyd (1993), and elsewhere. The door is thus open for some powerful and complementary investigations in which dynamical patterns act as the spatially and temporally extended vehicles of specific representational contents. Such developments, should they prove useful for understanding biological cognition, would not constitute a revision of the notion of internal representation itself (pace e.g. Van Gelder, 1995) so much as a revision of our ideas about the kinds of inner state and process that might act as the *vehicles* of such representation. The root notion of internal representation (weak-substantive or strong) remains unchanged, involving as it does only the idea of identifiable aspects of inner processing whose real functional role is to stand-in for other states of affairs (for an extended defense of this claim, see Miller & Freyd, 1993).

The putative change in our conception of likely representational vehicles is, however, profound and important. It constitutes a move away from the idea of static, text-like, spatially local, and atemporal vehicles to the much more challenging image of temporally and spatially extended patterns as the key players in a whole new kind of inner economy.

## 6. CONCLUSIONS: A SPACE FOR DEBATE

The dynamical challenge, I conclude, is substantive, important and all too easily misunderstood. On the vexed issue of internal representation, we can return a null-verdict. Given at least a weak-substantive notion of internal representation, it is an open empirical question whether the construct will earn its keep in explanations of biological cognition. A positive answer will require the isolation of distinct inner states or processes whose functional-adaptive role, as described in section 2, is to act as genuine stand-ins for extra-neural states of affairs. We saw that the contents of such putative representations need not be of the objectivist, observer-independent variety but could instead focus on functional-indexical, action-oriented features. Nonetheless, the project of isolating any such states is threatened

by the spectre of processes of continuous reciprocal causation criss-crossing organism, environment and inner neural sites. In this respect, strong internal representation may, somewhat surprisingly, fare better than weak since (by definition) strong representation supports episodes of environmentally de-coupled, off-line reasoning and is thus isolated from at least one source of causal complexity. The ubiquity of such off-line reasoning is, of course, open to question. The issues concerning representation thus reduce to questions about the isolability of inner content-bearing vehicles and the nature (weak or strong) of the standing-in relation itself. It was noted, however, that the tools of dynamical analysis may themselves provide a means of unpicking complex causal webs and revealing temporally and spatially extended entities as potential vehicles of representational content.

The issues about computation are even less clear-cut. There are powerful dynamicist considerations, at least in the area of motor control, that argue against any notion of complex, detailed, neural instruction sets and hence against a strong notion of stored inner programs. But we here confront not a dichotomy (programmed versus unprogrammed) but a continuum of possibilities linked to the amount and specificity of neural commands required to bring about a desired action. Moreover, the stored program idea may not even be essential to the more general image of the brain as some kind of computational device. Instead, all that may be required is the presence of semantically sensible transitions between representational states. Further resolution of these issues must thus wait upon much-needed progress in our general understanding of the nature of computation itself.

The true heart of the dynamical challenge, however, lies elsewhere. It lies in the vision of the brain as, when all is said and done, just another participant in the construction of situated action. From this vision proceeds an imperative: to give explanatory priority to the patterns that characterize embodied, situated action, to analyze those patterns (using the tools of collective variables, control parameters, coupled equations and the like), and against this background to seek further understanding of specifically internal contributions, architectures and organizations.<sup>10</sup> The real dynamical insight thus turns on issues of explanatory privilege (should we focus almost exclusively on the brain?) and temporality (should we focus on static states or on temporally extended processes?). And the advice is clear: look harder at temporally extended processes that span brain, body and world. This good counsel is the true fruit of the dynamical challenge.

**Acknowledgments:** This paper has benefited enormously from discussions with Randy Beer, Tim Van Gelder, Esther Thelen, Melanie Mitchell, Scott Kelso and Michael Wheeler. Thanks also to Francisco Varela, Esther Thelen, Bill Clancey and an anonymous referee for constructive and invaluable comments on an earlier version. I respectfully dedicate this paper to the memory of Donald Campbell.

## NOTES

1. See e.g. McClelland, Rumelhart and the PDP Research Group (1986) vols. I and II; Clark (1989); P.M. Churchland (1989); Clark (1993).

2. See e.g., P.M. Churchland (1989), McClelland et al (1986). Whether connectionists should have maintained this commitment is, of course, an open questions—see Wheeler (1994) and Clark & Toribio (1994) for some discussion.
3. This point is well-made by Beer (1995) who also argues that a failure to distinguish internal state and internal representation undermines the supposedly pro-representationalist observations of Kirsh (1991).
4. Thanks to Martin Davies for pointing this out.
5. These issues are pursued further in Clark & Grush (to appear).
6. See especially Grush (1995), Clark & Grush (to appear).
7. A lot more could (and should) be said about the BZ reaction (see e.g., Gyorgyi & Field, 1993, Thelen & Smith, 1994, Ch.3, and Rueger & Sharp, 1996).
8. Thanks to Esther Thelen & an anonymous referee for highlighting the importance of the continuity assumption.
9. Notice that this is by no means to trivialize the notion of computation. For most physical systems (e.g. the solar system) do not traffic in internal representations and hence will not count as computational. Nor will all neural activity count as computational: only that which is meant to mediate semantically sensible transitions between representational states (e.g. by transforming a representation of sentence structure from active to passive form, by transforming an agent-centered into an object-centered representation, etc).
10. Note that such a plea for explanatory priority does not imply that pure neuroscientific research should wait upon prior dynamical analysis. It serves rather to remind the neuroscientist that the brain is in the business of promoting successful action in a bodily and environmental context, and that nature is no respecter of a neat division of labor between brain, body and world.

## REFERENCES

- Amit, D. (1989). *Modeling brain function: The world of attractor neural networks*. Cambridge: Cambridge University Press.
- Ballard, D. (1991). Animate vision. *Artificial Intelligence*, 48, 57–86.
- Beer, R. (1995). A dynamical systems perspective on environment agent interactions. *Artificial Intelligence*, 72, 173–215.
- Beer, R. & Gallagher, J. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, 1, 91–122.
- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.
- Brooks, R. & Stein, L. (1993). *Building brains for bodies* (Memo 1439): Artificial Intelligence Laboratory, Cambridge, MA.
- Campbell, D. (1974). Evolutionary epistemology. in P. A. Schilpp (Ed), *The philosophy of Karl Popper*. La Salle: Open Court.
- Churchland, P. M. (1989). *The neurocomputational perspective*. Cambridge: MIT/Bradford Books.
- Churchland, P. & Sejnowski, T. (1992). *The computational brain*. Cambridge, MA: MIT Press.
- Clark, A. (1989). *Microcognition: Philosophy, cognitive science and parallel distributed processing*. Cambridge: MIT Press.
- Clark, A. (1993). *Associative engines: Connectionism, concepts and representational change*. Cambridge, MA: MIT Press.
- Clark, A. (1995). Moving minds: Situating content in the service of real-time success. In J. Tomberlin (Ed.), *Philosophical perspectives* (Vol. 10), pp. 89–104.
- Clark, A (1997) *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press.
- Clark, A. & Grush, R. (to appear). Towards a Cognitive Robotics. *Adaptive Behavior*
- Clark, A. & Toribio, J. (1994). *Doing without representing?* *Synthese*: 101, 401–431
- Dennett, D. (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster.
- Diamond (1990). Development and neural bases of AB and DR. In A. Diamond (Ed.), *The developmental and neural bases of higher cognitive functions* (pp. 267–317). New York: Nat. AC Sciences.
- Elman, J. (1991). *Incremental learning or the importance of starting small Technical Report 9101*. Center for Research in Language, University of California, San Diego.
- Fodor, J. (1975). *The language of thought*. New York: Crowell.
- Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge: MIT Press.

- Gleick, J. (1987). *Chaos*. New York: Viking.
- Grush, R. (1995). *Emulation and mind*. Unpublished MS.
- Gyorgyi, L. & Field, R. (1993). Modeling and interpretation of chaos in the Belousov-Zhabotinsky reaction. In R. Field & L. Gyorgyi (Eds.), *Chaos in chemistry and biochemistry* (pp. 47–85). Singapore: World Scientific.
- Harnad, S. (1994) (Ed). Special issue on What is computation? *Minds and Machines*, 4(4) 377–488.
- Haugeland, J. (1991). Representational genera. In W. Ramsey, S. Stich, & D. Rumelhart (Eds.), *Philosophy and connectionist theory*. New Jersey: Erlbaum.
- Heidegger (1961). *Being and time* (J. Macquarrie and E. Robinson, Trans.). (First published in 1927 ed.). New York: Harper and Row.
- Jordan, M., Flash, T., & Aron, Y. (1994). A model of the learning of arm trajectories from spatial deviations. *Journal of Cognitive Neuroscience*, 6(4), 359–376.
- Kirsh, D. (1991). Today the earwig, tomorrow man. *Artificial Intelligence*, 47, 161–184.
- Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.
- Maturana, & Varela, F. (1980). *Autopoiesis and cognition*. Dordrecht, Holland: Reidel.
- McClelland, J., Rumelhart, D., & Group, P. R. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. I & II). Cambridge: MIT Press.
- McNaughton, B. & Nadel, L. (1990). Hebb-Marr networks and the neurobiological representation of action in space. In M. Gluck and D. Rumelhart (Eds), *Neuroscience and connectionist theory* (pp. 1–64). Hillsdale, NJ: Erlbaum Associates.
- Merleau-Ponty, M. (1942). *La structure du comportement* (A. Fisher, Trans.). France: Presses Universitaires de France.
- Miller, G. & Freyd, J. (1993). *Dynamic mental representations of animate motion*. University of Sussex Cognitive Science Research Paper, 290.
- Mitchie, D. & Johnson, R. (1984). *The creative computer*. U.K.: Penguin.
- Nehmsow, U. & Smithers, T. (1991). Mapbuilding using self-organizing networks in “really useful colors.” In J. A. Meyer & S. Wilson (Eds.), *From animals to animats* (pp. 152–159). Cambridge: MIT Press.
- Petitot, J. (1985). *Les catastrophes de la parole*. Paris: Malone.
- Port, R., Cummins, F., & McCauley, J. (to appear). Naive time, temporal patterns and human audition. In R. Port & T. Van Gelder (Eds.), *Mind as motion*. Cambridge, MA: MIT Press.
- Port, R. & VanGelder, T. (Eds.) (1995). *Mind as motion*. Cambridge, MA: MIT Press.
- Ridley, M. (1985). *The problems of evolution*. Oxford University Press.
- Rueger, A. & Sharp, D. (1996). Simple theories of a messy world: Truth and explanatory power in non-linear dynamics. *British Journal for the Philosophy of Science*, 47(1), 93–112.
- Simonyi, R. et al. (1982). One dimensional dynamics in a multicomponent chemical reaction. *Physical Review Letters*, 49, 245–248.
- Skarda, C. & Freeman, W. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, 10, 161–195.
- Smithers, T. (1994). Why better robots make it harder. In D. Cliff, P. Husbands, J. A. Meyer, & S. Wilson (Eds), *From animals to animats 3* (p. 64–72). Cambridge, MA: MIT Press.
- Stein, B. & Meredith, M. (1994). *The merging of the senses*. Cambridge, MA: MIT Press.
- Thelen, E. & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Thelen, E. & Ulrich, B. (1991). *Hidden skills: A dynamic system analysis of treadmill stepping during the first year* (Serial No. 223, 56 (1)): Monographs of the Society for Research in Child Development.
- Van Essen, D. & Gallant, J. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, 13, 1–10.
- Van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy*, XCII(7), 345–381.
- Van Gelder, T., & Port, R. (1994). Beyond Symbolic: Towards a Karma-Sutra of Compositionality. In V. Honovar & L. Uhr (Eds.), *Symbol processing and connectionist network models in artificial intelligence and cognitive modeling: steps towards principled integration* (pp. 107–125). San Diego: Academic Press.
- Varela, F., Thompson, E. & Rosch, E. (1991). *The embodied mind*. Cambridge, MA: MIT Press.
- Wheeler, M. (1994). From activation to activity. *Artificial Intelligence and the Simulation of Behavior (AISB) Quarterly*, 87, 36–42.