

Explanation, Representation and the Dynamical Hypothesis

JOHN SYMONS

Department of Philosophy, The University of Texas at El Paso, El Paso, Texas 79968 USA; E-mail: jsymons@utep.edu

Abstract. This paper challenges arguments that systematic patterns of intelligent behavior license the claim that representations *must* play a role in the cognitive system analogous to that played by syntactical structures in a computer program. In place of traditional computational models, I argue that research inspired by Dynamical Systems theory can support an alternative view of representations. My suggestion is that we treat linguistic and representational structures as providing complex multi-dimensional targets for the development of individual brains. This approach acknowledges the indispensability of the intentional or representational idiom in psychological explanation without locating representations in the brains of intelligent agents.

Key words: dynamical systems theory, explanation, representation

1. Introduction

Many theorists who develop dynamical rather than sentential or propositional models of cognition deny that brains house content-bearing representational states. According to a growing number of neuroscientists (Skarda and Freeman, 1987) developmental psychologists (Thelen and Smith, 1994), roboticists (Brooks, 1991) and philosophers (Van Gelder, 1995), embodied cognition is not a matter of manipulating representations, or processing information in the brain. Instead, they argue, cognition is inextricably bound to action. The mind, they claim, is more like a hurricane or an epidemic than a flowchart or a logical proof. Hence, cognitive systems shouldn't be viewed as computer programs. Instead, they are a particular kind of dynamical system best studied using the mathematical tools of Dynamical Systems theory.

In response, mainstream philosophers of mind and many more traditional cognitive scientists contend that the dynamicists are far too radical. Their principal argument against the dynamical approach is that representations are an indispensable feature of any cognitive science worthy of the name. Explanations that fail to take the representational properties of the brain into account will fall prey to the same problems that undermined old-fashioned behaviorism.

This paper reviews the dispute between dynamicists and their opponents and outlines an approach that is roughly modeled on Dennett's intentional stance. The framework I propose, provides a way of understanding the crucial role of representation in psychological explanation that doesn't commit researchers to the principle that representations play a role in cognitive system analogous to that



played by syntactical structures in a computer program. Rather than viewing the brain itself as a representational structure of a certain kind, it may be fruitful to view the organism and the representational system as a co-evolving couple. We can continue to acknowledge that language-like structures, codes, representations, information processing models and the like, play a crucial heuristic role in the way we investigate the brain. However, rather than seeking the neural structures or processes that embody the tokens of particular units in a representational system, those representational systems themselves may be understood as a set of co-evolving environmental constraints on the development and activity of the brain. As a rough initial hypothesis, I suggest that language-like structures, as well as patterns of intelligent behavior can be understood as providing complex multi-dimensional targets for the organism. As we shall see, this approach can recognize the indispensability of the intentional or representational idiom in psychological explanation, without locating representations in the brains of intelligent agents.

Most of us already accept the notion that brains develop in order to meet the challenges of particular environmental niches (Edelman, 1987). So, it seems reasonable to treat the development of systematic patterns of intelligent behavior along similar lines. Clearly, individual human infants face the challenge of learning to navigate a rich symbolic and social landscape and in this sense it might seem obvious that baby brains develop to fit the language of the baby's social group. However, most of the tradition in cognitive science rests on Chomsky's assumption that such development would be impossible without a set of innate capacities, a preexisting grammar, or proto-language of some kind. This assumption, in turn, licensed the development of a cognitive science that saw systems of internal representations as its principal object of study. From the Chomskian perspective, human language acquisition looks like such an incredibly difficult task that it seems necessary to posit an innate crutch to help human babies participate in their language communities.

There is a viable alternative to the Chomskian picture of language acquisition. As Terry Deacon (1997) argues, explanation of language acquisition must work at two levels. On the one hand, individual babies certainly adjust to languages, and by itself this looks like a monumentally difficult task. However, over longer time scales, languages also adjust to the biology of human baby brains. Noting the power of Baldwinian evolution, Deacon and others have pointed out that the languages and conventional patterns of behavior that serve as the targets for the baby's development are both the sources of and *subject to* selective pressures.¹ If Deacon is correct, our innate capacities need not be the result of our possessing some set of internal representations or grammar, but can be explained instead in terms of the co-evolution of language and the brain. With Deacon's account, the need for innate language-like structures to handle the difficult task of learning language simply drops out. This is because the kinds of languages that survive through the generations are those that were easy enough for us to learn given the kinds of brains we possess. If we continue to suppose that there are innate

structures (grammars) on Deacon's model, then they are an emergent product of a co-evolutionary relationship between language and the brain rather than freakishly anomalous devices in the brain that recursively generate grammatical utterances. From Deacon's perspective our linguistic capacities are best studied by attending to the properties of dynamical and co-evolutionary systems rather than by positing a set of internal representations or mechanisms in the brains of individual speakers.

If the Chomskian demand for internal representational systems can be deflated along the lines suggested by Deacon and others, then perhaps it is possible to view a great deal of our mental lives as an essentially social art. Much, perhaps all, of what we consider intelligence, might be most appropriately studied as part of an active, often socially-mediated relationship between organisms and their environments. On this view, representations are part of a social practice that permits us to track the behavior of massively complex biological systems in an enormous variety of contexts. As such, the notion of representation has its primary home in the practice of predicting and explaining behavior. Given such an approach, representations can still be said to 'cause' behavior, albeit, insofar as the brains of agents have developed to negotiate the patterns of behavior and the complex symbolic and representational systems that emerge in social contexts. If representations (traditionally construed) can be said to exist, they will exist as emergent phenomena in social groups rather than as components in the mechanisms of the brain.

Accounting for the ubiquity and importance of representations in psychological explanations in the way I propose is likely to be labeled *instrumentalist*. However, it can be argued that the representation-like phenomena that play a central role in ordinary psychological explanations may still be said to have at least as much reality as they have in any sample, hard-core realist theory of internal representation. Calling a particular pattern of brain activity a representation is no more *realistic* than treating representational structures as targets for the development and action of the brain. Representations are real enough for the dynamicist, they're just not in the brain.

Section 2 presents a brief outline of the dynamicist position before turning in greater detail to the relationship between explanation and representation in traditional cognitive science. Sections 3 and 4 treat objections to the dynamical hypothesis based on the indispensability of representation. In the following sections, the notion of representations as inner causes is criticized along the lines suggested by Dennett and others. The paper concludes by suggesting an alternative role for representation that avoids the biologically implausible tenets of the computational model of representations as inner causes of intelligent behavior.

2. The Dynamical Hypothesis

Dynamical Systems theory provides a useful set of mathematical tools for the study of cognition in biological settings. During the 1990's researchers working with Dynamical Systems theory began to advocate a view of intelligent behavior that they

characterized as fundamentally opposed to the basic tenets of traditional cognitive science. In a series of articles and books, dynamicists vigorously challenged what they called the *computational approach* to mental life (see especially Port and Van Gelder, 1996). While dynamicist characterizations of the computational paradigm have been faulted for a lack of precision (see, e.g., Eliasmith, 1998) the principal target of their criticism is clear. According to the dynamicists, traditional cognitive science is governed by the mistaken assumption that the biological mechanisms underlying cognition are best studied as the sequential manipulation of discrete symbolic structures. The computational approach rests on transformations over stable symbolic structures in the brain, dynamicists argue that this analogy between the way human brains work and the way digital computers work has obscured important dynamical properties of biologically-based cognitive systems. From the perspective of the dynamicists, the function and physiology of the brain is so radically unlike the idealized digital computers and neural networks of traditional cognitive science as to render the computational approach an implausible model of cognition in biological systems in principle.

In an effort to remain faithful to the biological, behavioral and temporal phenomena associated with cognition, dynamicists offer mathematical models of spatio-temporal processes in the brain and behavior that differ in a variety of important ways from traditional efforts in cognitive science. As mentioned above, the most philosophically interesting difference between dynamical models and most traditional cognitive science is that the dynamicist, unlike the cognitive psychologist or the software designer, doesn't look for explanations of cognitive phenomena through an account of the etiological relationships between symbolic structures. In place of models that employ symbolic structures governed by discrete mathematics of one kind or another, dynamicists advocate mathematical models of biological and behavioral patterns that mirror the continuous nature of biological phenomena. In effect, the principal difference between the dynamicists and the computationalists is that dynamical models range over spatio-temporal structures rather than quasi-linguistic declarative structures. A dynamicist might, for example, use differential equations to capture the way a system is changing at a particular time as a function of its state at that moment. More traditional computational models are usually understood as being governed by transformation rules from one set of symbols or representations to another.

The mathematical tools that have encouraged the increasingly critical attitude towards computationalism belong, first and foremost, to the study of a set of *mathematical* objects known as dynamical systems. The variables used to define these systems can be thought of as the dimensions in a multidimensional space. This is what dynamicists mean by the *state space* for the system. Given this stipulation, the state space will, by definition, include the set of all possible states of the system in question. We can then track the changing states of an observed system whose behavior we are modeling mathematically, as a trajectory through this space. Interesting systems often show regular patterns in their trajectories through their state spaces.

These patterns represent what scientists call the attractive properties of the state space in question. A system's trajectory may tend towards one or many points or cyclical patterns. The attractive properties of the state space for a system can be represented topologically in order to graphically reveal the paths and points towards which the system's trajectory is likely to tend given varying initial conditions.

Dynamical Systems theory allows us to model biological phenomena as an enormous set of states of a system and the mapping of those states onto successors. While these may, in principle, be continuous or discrete in their description, dynamicists with an interest in cognition emphasize continuity.² Dynamical systems can be conservative, where the states map on-to-one (phase space conserving), or dissipative, where the mapping may be many to one. Conservative systems can be thought of as isolated systems, systems that don't exchange energy with their environments (however, they may have elastic collisions in which momentum but not energy is exchanged).³ Dissipative systems can be thought of as systems that can lose energy to their environment.⁴

For Dynamical Systems theorists, accounts of the behavior of these systems drawn from these models can constitute useful explanations. This is the point where the dynamicists and the traditional computationalists part ways. Clearly, the shift in emphasis from the study of symbolic transformations to the study of the spatio-temporal dynamics of the brain clearly involves more than the use of a new set of tools. There is a basic difference between dynamicists and their opponents over what it is for something to count as an explanation in cognitive science. As Bechtel (1998) and Van Gelder (1995) have noted, this difference centers on the role of representation in the explanation of cognitive phenomena. Critics of the dynamicist movement have correctly emphasized the importance of representation in the explanation of any genuinely cognitive phenomenon, whereas advocates like Van Gelder question the very idea of saying that biological systems work via a system of commands or messages.

The growing popularity of DST in cognitive science has encouraged some philosophers to argue for what Tim van Gelder has called *the dynamical hypothesis*. The dynamical hypothesis makes both ontological and epistemological claims about mental life – about both what the mind is: *the nature hypothesis*, and how we can best come to know the mind: *the knowledge hypothesis*. (See van Gelder (1999) and for critical discussion Chemero (2000)). While the nature hypothesis is the claim that cognitive agents *are* dynamical systems, the knowledge hypothesis is the epistemological claim that cognitive agents *should* be investigated using the tools of DST. Both aspects of the dynamical hypothesis have struck some readers as trivial. For example, critics like Eliasmith (1996) note that simply calling cognitive systems dynamical (or claiming that they can be understood using the tools of dynamical systems theory) is compatible with almost every traditional theory in cognitive science. Without further restrictions on what can and cannot count as a dynamical system, even the humble Turing machine can be construed

as a dynamical system.⁵ And so, at least at the outset, it might seem that there is nothing revolutionary entailed by the hypothesis.

Of course, van Gelder and others attach a far stronger set of implications to the dynamical hypothesis. They see the spirit of the hypothesis as presenting a genuine challenge to the computationalist orthodoxy. They suggest, for example, that “dynamics provides not just a set of mathematical tools but a deeply different perspective on the overall nature of cognitive systems. Dynamicists from diverse areas of cognitive science share more than a mathematical language; they have a common worldview” (Port and Van Gelder 1995, vii). This worldview is marked first and foremost by a critical attitude towards computationalism; by the belief that cognition does not involve representation and that standard computational models are not up to the task of explaining mental life. In opposition to the computational approach to mental life, dynamicists envision the dynamical approach as “a fully fledged research program standing as an alternative to the computational approach” (ibid. vii). The principal difference between dynamicists and computationalists concerns the role of representation in their explanations of cognitive phenomena. As we shall see in the next section, traditional cognitive science and philosophy of mind has placed representation at the heart of the picture of explanation, whereas the dynamicists see patterns of behavior and action as central to the explanation of cognitive systems.

Once they abandon the goal of explaining cognitive phenomena in terms of “internal representations” (Van Gelder, 1995, 346), dynamicists must face the fact that explanations involving representations have seemed to work very well in cognitive science. Doing without representations that serve the role of inner causes of behavior means abandoning much of the most interesting and important work done in cognitive science. In order to justify this step, dynamicists must offer an alternative explanatory framework to a view that regards representations as innercauses of behavior.

3. Representation

From the perspective of most cognitive scientists and philosophers of mind, any legitimate study of cognition must acknowledge the mind’s representational nature. Explanations in cognitive science have been understood to differ from explanations in biology, physics or chemistry insofar as most cognitive scientists (at least traditionally) have held that genuine explanations of cognitive phenomena, unlike explanations in the non-cognitive sciences, *must* involve semantically-evaluable *and* etiologically-involved entities (Fodor, 1987).

So, for example, in his *Mind: An Introduction to Cognitive Science* Paul Thagard characterizes the basic schema for explanations in cognitive science as follows:

Explanation Target

Why do people have a particular kind of **intelligent behavior**?

Explanatory pattern

People have mental **representations**. People have algorithmic **processes** that operate on those **representations**. These **processes**, applied to the **representations** produce the behavior.

(1997, 19 (author's emphasis))

Thagard's characterization of explanation in cognitive science is deliberately ecumenical with respect to the kinds of things that can count as *representations* or *processes*. As he suggests, his definition is broad enough to include a considerable portion of what was once considered *nonclassical* connectionist work within the purview of cognitive science (Thagard, 1997, 22). While connectionism is opposed to many of the tenets of traditional computationalism regarding the nature of representations in cognitive systems, both views are united in regarding representation as the heart of explanation in cognitive science. For all its inclusiveness, Thagard's schema demonstrates the core commitment of traditional cognitive scientists. For Thagard, like most classical cognitive scientists, thinking is best understood in terms of processes that operate on representational structures in the mind, whether they regard representation as distributed or nodular is an internecine dispute. Dynamicists like Van Gelder are working against the central tenet of what Thagard calls 'The Computational Representational Understanding of Mind' (CRUM) and, needless to say, they have encountered significant resistance in the philosophical community.

At this point, applications of Dynamical Systems theory lack the kind of neat theoretical framework that philosophers and psychologists have found in CRUM. Insofar as they eschew representational structures, advocates of dynamics have been criticized for attempting to drag psychology back to the dark days of behaviorism. However, for dynamicists, CRUM's focus on rules and representations has misled cognitive scientists and has obscured the essentially dynamical nature of cognition. Dynamicists believe that, if there are representations, then they are unlikely to play the kind of role in the brain that representations play in the structure of a computational system. In characterizing the computationalist position, Port and Van Gelder catalog five mistaken assumptions underlying the computationalist perspective:

1. Representations are static structures of discrete symbols.
 2. Cognitive operations are transformations from one static symbol structure to the next.
 3. These transformations are discrete, effectively instantaneous, and sequential.
 4. The mental computer is broken down into a number of modules responsible for different symbol-processing tasks.
 5. At the periphery of the system are input and output transducers: systems which convert sensory stimulation to input representations and systems which convert output representations into physical movement
- (see Port and Van Gelder, 1995, 1)

While it is possible to find occasional examples of traditional work in the computationalist tradition that diverge from these principles, Van Gelder and Port are criticizing the general tendency in cognitive science to see cognition as a symbol-manipulating process that can somehow be understood apart from particular biological instantiations and which is essentially insulated from perception and action. In place of the static and insulated picture of cognition, they see cognitive systems as “a structure of mutually and simultaneously influencing change. Its processes do not take place in the arbitrary discrete time of computer steps; rather, they unfold in the *real* time of ongoing change in the environment, the body, and the nervous system. The cognitive system does not interact with other aspects of the world by passing messages or commands; rather, it continuously coevolves with them” (1995, 3). For the dynamicists, the computational model of mind is an idealization that distorts the continuous and dynamical properties of real biological cognition.

In terms of what a biological system is actually doing when it's thinking, there can be little doubt that the dynamicists offer descriptions that are more accurate, or at least more plausible than the computationalists. However, in terms of the *explanation* of cognitive systems, computationalists, with their emphasis on compositionality and intentionality seem to have the upper hand. The following section focuses on the problematic status of dynamical accounts of cognition *qua* explanation.

4. The Uniqueness of Cognition

Two kinds of objections have been raised against models employing Dynamical Systems theory. On the one hand, some philosophers have doubted whether these researchers can really live up to their own hype and have urged us to rethink the methodological fruitfulness of the approach. While many of these methodological criticisms of applied Dynamical Systems theory have some merit, they are not the focus of the present paper. A second, and more philosophically interesting, kind of criticism stems from the ineliminability of certain systematic characteristics of cognitive phenomena (satisfaction conditions, compositionality, intentionality) and the claim that these characteristics cannot be captured by a non-sentential kinematics of the mind and brain.

Critics agree that explanations employing Dynamical Systems theory can meet some very general and uncontroversial requirements for being considered good scientific explanations. For example, three such conditions that any scientific explanation should meet are truthfulness, the ability to illuminate counterfactuals and the tendency to increase the unity of our worldview. Since dynamical models predict both actual and potential changes of the systems they study, they easily satisfy the requirement that an explanation should shed light on counterfactuals.⁶ Furthermore, we can assume that many applications of Dynamical Systems theory generate predictions and explanations that are, at least approximately, true.⁷ Finally, such theories clearly pick out certain universal properties in real dynamical

systems, thereby contributing to the overall unity of our scientific picture of things. The role of dynamics as a unifying, all-purpose framework for treating things as diverse as hurricanes, epidemics and traffic jams has led critics to suspect that it fails to capture the uniquely cognitive aspects of intelligent behavior. Dynamics seems to apply to the behavior of systems of all kinds, whether or not they happen to be thinking things. Van Gelder has responded to this charge by suggesting that we should see cognitive agents as members of a subset of dynamical systems with particular kinds of behaviors that distinguish us from hurricanes and traffic jams.

understanding cognitive agents as dynamical systems 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).

A more dramatic example, which might serve to make Van Gelder's point is Walter Freeman's famous model of the olfactory bulb in rabbits. Without getting into too many of the details, Freeman's goal is to explain the emergence of intentionality in the rabbit's olfactory system by presenting a system that can alternate between chaotic activity corresponding to the learning phase and orderly trajectories corresponding to specific scents (Skarda and Freeman, 1987). Contrary to the anti-dynamicist argument from the uniqueness of cognitive or intentional systems, Freeman's account of chaotic brain activity explicitly pinpoints four properties of the adaptive capacity of cognitive systems that he claims distinguish them from non-cognitive systems. These properties capture what Freeman believes to be the essence of intentional behavior. By modeling the unique adaptive capacities of the olfactory bulb without recourse to transformations over symbols, Freeman's model seems to provide a biologically plausible alternative to computational treatments of the sensory systems. Freeman has suggested that his non-representational characterization of the olfactory system can be generalized to brain function more generally.

Of course, Freeman could be wrong about the four preferred properties that he identifies with intentionality, or he could be wrong about the application of his model to the actual behavior of the olfactory system. However, the problem for most philosophers of mind is not whether he is right or wrong in empirical detail, but whether it makes any sense to call a system with no representations a cognitive system at all?

According to Jerry Fodor, for example, the systematicity of intelligent behavior is proof that it *must* rest on a language-like structure of representations.⁸ Such representations play an incidental (if any) role in the dynamical models we are considering. Therefore, according to critics, dynamical models will, ultimately

have as little relevance for cognitive science as behaviorism had for linguistics after Chomsky. Such arguments against dynamical models arise from the idea that psychological explanations must account for two basic properties of intelligent behavior:

1. Intelligent behavior is compatible with intentional explanation. Hence any general theory of cognitive systems should permit a notion of *content* that is adequate to account for the success of intentional or *folk-psychological* (belief – desire – action) models of explanation.
2. Human linguistic competence exhibits compositionality. The satisfaction conditions for complex linguistic expressions are determined by the satisfaction conditions of their grammatical parts. Likewise a semantics for propositional attitudes should make clear how the satisfaction conditions of beliefs desires and the like are determined by those of their constituent concepts, thereby explaining why minds are typically productive and systematic (from Fodor, 1997).

For the classical tradition in cognitive science, the semantic coherence of human thought and the systematic and generative power of human grammar seemed to provide sufficient proof that explanations of cognitive phenomena will have to do more than merely *describe* the complex interaction of a mass of neural activity in the brain. Instead, there will be a matter of discovering the rules governing the combinations and transformation of physical symbols. Hence, according to Fodor for example, our best bet is to suppose that thoughts are composed of other thoughts in a language-like way. The explanation of the rules governing composition is itself equivalent to a computational structure (a device or grammar of some sort) for generating the infinite number of grammatical sentences in our language.

When Chomsky puzzled over the seemingly miraculous ability of children to become fluent speakers of their language, he came to the conclusion that the systematic properties of language were conclusive evidence for the existence of internal mechanisms ranging over and generating representations. Babies could never learn to understand language or to produce novel and grammatical sentences without some kind of innate mechanism that disposed them to do so. As we saw in the introduction, the kinds of coevolutionary models of language acquisition found for example in Deacon's work provide a possible answer to Chomskian objections (see Symons, forthcoming). However, even if we could develop an alternative model of language acquisition to Chomskian linguistics, we would be left with the still deeper problem of accounting for the efficacy of our commonsense intentional psychology.

It is often argued that the truth of our commonsense psychology depends on the reality of syntactically structured representations as the inner causes of intelligent behavior. This view seemed to gain considerable support during the heyday of the artificial intelligence movement. The problem for such views is that these days few researchers believe that brains bear any theoretically useful resemblance to digital computers. Nevertheless, because of the explanation argument, most work

in cognitive science continues to be tied to the idea that the computer program, the pattern of syntactically structured representations serving as the inner causes of intelligent behavior is an appropriate model for understanding cognition.

So, if the brain itself does not work like a digital computer, does this mean that the cognitive scientist's notion of representation is devoid of empirical content? Not necessarily. Philosophers have pointed out that the notion of representation plays a crucial role in grounding the generalizations of psychological theory. In fact, it has been difficult to imagine a well-developed psychological theory that could do without the assumption that minds represent the world in some way. The crucial importance of representational and intentional notions in psychological explanation that has encouraged most philosophers of mind to see traditional rules-and-representations style computational cognitive science as the only game in town.

Philosophers like Fodor (1990, 156) and Putnam (1988) have treated attempts to do away with intentional notions as self-defeating attacks on the core of human knowledge. In one famous passage, Fodor claims that if talk of representations were to collapse "that would be, beyond comparison, the greatest intellectual catastrophe in the history of our species" (1987, xii cited in Stich, 1996, 169). Resting the viability of common sense intentional psychology on the existence of entities in the brain that are both semantically evaluable and etiologically involved strikes me as a risky strategy. Fodor might (and Putnam almost certainly would) argue that he isn't really resting the truthfulness of folk psychology on the existence of causally efficacious representations in the brain. However, the dire warnings that Fodor and others regularly offer in the face of models of the brain and behavior that don't involve representations tell a different story.

While Fodor represents the extreme end of what Dennett has called the 'hysterical realist' attitude towards representations, he is certainly entitled to his concern for the integrity of commonsense intentional psychology. However, it's a mistake to argue that folk psychology totters simply because the brain doesn't traffic in content-bearing representations. The dynamicists are not arguing that our ordinary psychological explanations are always simply false. For example, it would be folly to argue that it's always untrue to say that:

'Charles brought his umbrella because he thought it was going to rain'

or that

'He won't order the refried beans because he doesn't want to eat lard'.

Philosophers are correctly wary of theories that lead us to say that these kinds of folk psychological statements are necessarily false. When Fodor and others warn of the dangers of Churchland-style eliminativism or Dennettian irrealism, their warnings should be taken with a grain of salt.⁹ This kind of rearguard action in defense of the computational-representational model of mind relies, for the most part, on scare tactics. As we shall see, the dynamical approach to cognitive science leaves folk psychology safe and sound.

5. Some Difficulties with the ‘Representations as Inner Causes’ Thesis

The most prominent inner-cause theory in the philosophy of mind is generally associated with Fodor’s ‘language-of-thought’ hypothesis (see Fodor, 1975). While this view of the mind has had considerable influence in cognitive psychology (see e.g. Garfield, 1989), it rests on the questionable philosophical presupposition that psychological generalizations, as well as particular true propositions in psychology, must make reference to representations that function as inner causes in order for those generalizations and propositions to be true. On this view, intelligent cognition involves mental representations with decomposable structures of a variety of kinds, while cognition itself is a matter of the causal interaction of these structures. Consequently, cognition can be understood in terms of the algorithms that govern the causal relations among these structures.

So, for philosophers like Fodor, if it is true that Jane voted for Nader because she believed him to be the best candidate then it must also be true that a belief, ‘Nader is the best candidate’ *caused* her action. Fodor warns of dire consequences if this turns out not to be true. The truth of folk-psychological assertions like this, according to Fodor, implies the truth of the ‘language of thought’ hypothesis. And if this hypothesis is correct, then the brain, which controls Jane’s bodily movements, is governed by processes that are, in turn, governed by language-like (syntactically structured) representations. Fodor argues that the belief must exist in Jane’s brain in such a way that it can simultaneously cause the actions that constitute voting *and* interact with other representations in a reasonable way. Her beliefs must be organized and interact in the same way in the brain as they do in our commonsense reasoning. The representation of her belief ‘Nader is the best candidate’ must exist in Jane’s brain in such a way as to make it compatible with certain other representations such as ‘Nader is a candidate,’ ‘There can be only one best candidate,’ etc. Furthermore, it must exist in such a way as to make it incompatible with the belief that Bush or Gore are better candidates than Nader.

For the ‘language of thought’ hypothesis to have any empirical content, our ordinary folk-psychological generalizations and habits must not merely provide a useful way to predict the behavior of our fellows, but must also offer a source of scientific insight into the inner workings of the mind/brain. The basic assumption underlying Fodor’s philosophical reflection is the belief that, in order for folk psychological statements to be true, they must be statements about the inner causes of our actions. At this point, dynamicists should argue that it is an unwarranted philosophical step to go from the truthfulness of folk-psychological statements to the idea that those statements are uncovering some kind of inner mechanism.

The inner cause thesis has a number of important weaknesses that make it very unlikely that it will be a lasting part of the brain and behavioral sciences. So for example, it is extremely improbable that neuroscience will uncover distinguishable mechanisms that would be possessed by everyone sharing a particular belief. We are unlikely to find the kind of internal mechanism that could be at work in *all*

instances of such phenomena as wanting a sandwich or believing that Australia is surrounded by water. There are enormous barriers to the kinds of future discoveries that the inner-cause thesis presupposes. Recognizing this should lead us to doubt whether we will ever discover the kinds of representations in the brain that Fodor assumes must exist.

But if this is so, then how are we to ground the truth of our folk-psychological claims? We use the notion of representation with considerable success when talking about everything from chess computers to birds to people. While this strategy may be useful and may even generate objectively true statements, the dynamicist can argue that we are not thereby licensed to infer that the structure of our ordinary psychological generalizations can serve as a model for the mechanisms at work in the human brain. Take for example the following statements that include representational or intentional notions:

- (a) The only reason Fido is obeying you is because he *wants* a biscuit.
- (b) I won't move my pawn because I can see that the computer *wants* me to leave my rook unprotected.
- (c) She buys lottery tickets because she *thinks* she has a chance to win.

What makes statements like (a) – (c) true or false? Is it true that the chess computer *wants* me to leave my pawn unprotected? In some obvious sense, the answer is yes. However, one might doubt whether the chess computer wants to win in the same way that the dog wants a biscuit. And surely the computer and the dog *want* in a way that differs again from the way the person in (c) *wants* to win the lottery.

While traditional philosophers might urge us to dig deeper into the minds of the creatures in (a) – (c), as we shall see, the appropriate strategy is to look first to the behavior of the person characterizing the cognitive systems in each case. What (a) - (c) share in common is *not* that they all refer to some fundamental internal state corresponding to desire. It is highly unlikely that an examination of the innards of the three *wanters* would reveal some physical structure corresponding to their wants that they shared in common. Furthermore, even if there were some physical structure shared in common, it is almost certain that we could find a fourth *wanter* who wouldn't have it.

So, although we should not necessarily exclude the existence of neural patterns or structures that are common to animals that 'want', these structures are generally irrelevant to the kinds of claims we make about the thoughts and desires of our fellows. Instead, (a) - (c) are all instances of reports from what Dennett calls the *intentional stance*, they are all shorthand accounts of the predictive strategies that their speakers and hearers have adopted towards the world

For Dennett, one can be right about someone's beliefs or desires in the same way one can be right about whether someone would make a good husband or a good mayor. For example, I believe that my friend Marie would make a great mayor. In my judgment, her character and abilities suit her for the position and bode well for great success in public office. However, saying this does not imply the existence

of something like a good mayor gene in her cells or an essence of *good mayorness* in her soul. I am simply predicting that, given the right opportunities and the right circumstances, Marie would perform very well as mayor. Similarly, when we say that someone or something wants or believes something, the truth of our claim is not a matter of correctly locating, or picking out some internal structure in the person's brain or mind. Instead, our ascriptions of belief and desire are, for the most part, reports on the strategy we are using to predict that person's behavior. If we have adopted a good strategy for predicting the behavior of the person or animal in question, then we have at least some reason to call our statements about that person's state true.

For those who reject Dennett's approach and insist on the necessity of representation in psychological explanations, the unavoidable problem will be the difficulty of connecting the intentional entities that psychological explanations seem to call for, with the messy dynamical details of neuroscience. If one assumes that that certain generalizations of psychological theory or, for that matter of ordinary folk-psychological explanation, are true because they reflect an underlying structure with real causal efficacy, then psychological explanations must make contact with processes in the brain. So, while most philosophers have agreed with Fodor's claim that the notion of representation plays an essential role in psychological explanation, far fewer follow him when he argues that the truth of psychological explanations depends on the reality of an underlying language of thought. If we take the language of thought hypothesis (LOT) seriously, i.e., if we believe that it has empirical content, then it entails that processes in the brain are causally governed by semantically evaluable patterns that, in turn, are structured by the syntax of a universal language of human thought. In order for LOT to be true, dynamical processes in the brain would have to be quite different from those we actually find. As Garson (1995) and Dennett (1977) have noted, taking biologically plausible models of cognitive systems seriously means that you'll have a hard time defending an empirically meaningful version of LOT.¹⁰

Furthermore, the argument that the systematic properties and the compositional character of intelligent behavior *necessitate* the existence of an LOT style algorithmic structure, is simply incorrect. By now, there are many examples of mechanisms that dynamically produce the kind of recursive embedding that language-like compositional structures require *without* having the desired syntactical and semantical rules included in the system from the outset. For example, in one of the earliest philosophical papers on the dynamical approach, Horgan and Tienson (1994) described George Berg's (1992) connectionist model of natural-language sentence parsing. Such systems accomplish their task without recourse to algorithms over representations.

6. How can we Manage without Assuming that Representations are Inner Causes?

The virtues traditionally attached to psychological explanations that involve representations, do not force us to confine ourselves to a methodology that infers the existence of representations in the brains (or minds) of the subject whose behavior is being explained. At this point in the history of cognitive science there is no shortage of models that perform sophisticated tasks, including tasks that we would traditionally call information-processing in ways that do not involve the implementation of algorithms ranging over representations.

Furthermore, the claim that positing representational structure is an indispensable "guide to science" (see for example Chemero, 2000) is undermined by the fact that many fertile explanations draw, for example, on the attractor dynamics of a system as a guide to further investigation of high-level cognitive capacities. So, for instance in their study of word recognition, McLeod, Shallice, and Plaut (2000) demonstrate the key role of attractor structure in the ability of certain recurrent connectionist networks, normal subjects and dyslexic patients to map orthography onto semantics. The converging evidence from these three groups turns out to be the attractor structure that each seems to manifest. This allows McLeod et al. to suggest that the study of attractor structure within cognitive psychology holds considerable explanatory power.

Nevertheless, the basic problem with non-representational models is philosophical rather than methodological or practical. The bottom line is that the dynamicist movement is faced with the difficulty of accounting for the relationship between the models that they propose and the conventional folk-psychological patterns of explanation that serve us so well. In response to this difficulty, connectionists and others have sought biologically plausible ways of accounting for the power of the mind to represent.

Representations have been thought of as distributed throughout a connectionist network or as emergent patterns of one kind or another in these systems. Andy Clark, for example, has attempted to reconcile dynamics and computation through his notion of action-oriented representations in brains and connectionist networks. (see Clark, 1998) Connectionists who interpret their machines as systems of distributed representations are pulled in two opposing directions. They attempt to retain representations as an indispensable component of our explanatory framework while simultaneously working to generate a biologically plausible way of instantiating representations in the brain (Horgan and Tienson, 1996). Unfortunately, such ecumenical strategies inherit the weaknesses of both the dynamicist and the computationalist approach to representations. On the one hand, critics like Jerry Fodor argue that the kinds of representations connectionists identify, fail to play the kind of compositional and intentional role in psychological explanation that had seemed indispensable to psychological explanation in the first place (Fodor and Pylyshyn, 1988; Fodor and McLaughlin, 1990). And on the other hand

neuroscientists like Walter Freeman have pointed out some of the biologically implausible properties of some of the most prominent connectionist networks (1988, 1990, 1991).

While proponents of dynamical systems have biological plausibility on their side, they must provide a way to explain the unusual fecundity of *psychological* explanations featuring representations. In ordinary folk-psychological explanation as well as in more formal computational cognitive science, explanations involving representations have a level of generality and usefulness that dynamicist models have yet to match. If they are to respond to philosophical critics, dynamicists must account for the usefulness of the notion of representation while acknowledging the dynamical nature of cognitive systems in a biological context.

Fortunately, there is a readily available theoretical framework for integrating work in the dynamical tradition with folk-psychological uses of representation. According to Dennett, ascribing a 'belief' or a 'desire' to some system is not the same as describing some portion of an animal's brain. Instead, our use of notions like belief and desire, according to Dennett is more like performing a mathematical calculation than describing a mechanism or process in the animal's brain.

For Dennett, beliefs and desires are like centers of gravity, equators or lines of force – they are virtual tools that allow us to simplify the behavior of otherwise massively complex systems. For example, an astronomer might, for the sake of simplicity, treat the motion of a planet in terms of the motion of a single point, a center of gravity. The astronomer would almost certainly agree that the center of gravity isn't *really* hidden deep in the planet's core. He or she would tell us that the center of gravity is a theoretical device that allows us to conveniently track gravitational systems with the minimum of extraneous calculation. In a sense, the center of gravity is something that exists only in the calculations of the astronomer. This is the case even though the predictions that the astronomer's calculations provide might be extremely accurate. It's not that Dennett is denying the objectivity of statements about belief, rather he is committed to denying that beliefs, desires and other intentional entities exist in the same way that the entities of the physical science exist.

Critics have taken Dennett's position to imply that mental life doesn't really exist, and that folk psychology is merely an illusion. This charge is associated with the labels *interpretationist*, *relativist* and *instrumentalist*. Many philosophers have argued that Dennett's instrumentalist position with respect to the mind makes little or no sense since it denies the reality of something we know with utmost certainty. It seems insane to deny that beliefs are really *in* the mind or brain of the thinker. Similarly, it seems to make little sense to say that the only reason you or I have beliefs or desires is because some third person interprets us as having them. If it were solely a matter of interpretation, then that third person would himself be an interpreter only in the eyes of some fourth interpreter, who would in turn be an interpreter in the eyes of some fifth interpreter, and so on into an infinite regress. Arguments such as these have convinced many that Dennett must be wrong.

However, Dennett should not be understood as denying the truth of statements about belief. For example, it is true that as I write these words, alone at my desk, I believe my coffee is cold. This is true, despite the fact that there is nobody here to interpret my behavior. In an effort to clarify his position, Dennett's 1991 article, 'Real Patterns' reminds us that the virtual objects he describes, like centers of gravity and planetary equators are more than mere arbitrary interpretations of reality. They allow us to understand and predict *real patterns* of behavior. Ultimately, Dennett believes that the truth of an ascription of belief or desire to another biological or mechanical system is determined by whether the ascription would allow one to generate accurate predictions. Dennett summarizes his attitude towards the existence of beliefs as follows:

My thesis is that while belief is a perfectly objective phenomenon (which apparently makes me a realist), it can be discerned only from the point of view of someone who adopts a certain predictive strategy, *the intentional stance* (which apparently makes me an interpretationist). (1988, 496)

Dennett believes that it takes at least two animals in an ecological setting where behavior must be reliably anticipated before we can even begin talking about mental life. Whereas traditional philosophers and psychologists have treated our ordinary talk of beliefs and desires as a kind of proto-scientific theory about the contents of a private inner world, Dennett's approach focuses on the practical role played by those terms from the perspective of the objective observer. As he shows throughout his writings, the virtual world of belief and desire only emerges once animals are in the business of predicting and interpreting one another.

Like Dennett, Donald Davidson has argued that a necessary condition for the emergence of thought, is the configuration of at least three elements: two creatures who can interact with one another, and a shared set of stimuli. "Without the triangle there are two aspects of thought for which we cannot account. These two aspects are: the objectivity of thought, and the empirical content of thoughts about the external world" (12). The triangle that Davidson describes, stands for the simplest possible interpersonal context. Once this is in place, Davidson can have those aspects of thought that were impossible to detect in a purely physicalist description of the world, namely, the possibility of error and the possibility that experience gives some content to our thought. Of course, this is by no means sufficient for the emergence of thought, but by establishing this configuration as a necessary condition for the emergence of the kinds of properties that Davidson associates with thought, he shows that the conceptual difference between the thinking and unthinking context is due to the novel interpersonal configuration, rather than any physical structure or set of dispositions in the creatures before they interacted. It is true that Davidson's triangle is a chance affair from the perspective of some lower level science. It is also true that the kind of configuration that Davidson and Dennett both (in significantly different ways) rely on can only be considered a salient explanation of the emergent phenomena once folk psychology is already well under way. However, both potential objections apply to all non-physical sciences. The

kinds of folk-psychological regularities that unfold when Davidson's triangulation of organisms and environment is in place is certainly a mystery from the perspective of physics and in that sense, the lower-level science cannot completely explain the emergence of the higher-level phenomena. But, even though the explanation is offered in retrospect, it nonetheless relies only on naturalistic conditions.

Davidson's triangulation model of the emergence of thought is admittedly a toy example, a radically simplified form of the kind of solution that the problem of representation requires. However, we can see scientific applications of this approach to the emergence of thought in a more sophisticated manner in the co-evolutionary accounts of the origin of language and thought presented by scientists like Deacon. From a scientific perspective, it is clear that the kind of story that needs to be told about the emergence of thought will require attention to the interplay of multiple factors, including evolutionary, neuroscientific and social constraints.

So, what is the upshot of the position taken by Dennett, Davidson and others for the dynamical approach? Given our current knowledge of neural plasticity, it is quite reasonable to assume that the brain is a dynamical system that follows a strategy of self-modification in response to the kinds of systematic properties that intelligent behavior manifests. Understanding representations as patterns in a social context is certainly not a new idea. For example, Quine's famous topiary analogy in *Word and Object* beautifully illustrates his commitment to understanding language as a "social art" (1960, ix):

Different persons growing up in the same language are like different bushes trimmed and trained to take the shape of identical elephants. The anatomical details of twigs and branches will fulfill the elephantine form differently from bush to bush, but the overall outward results are alike. (1960, 8)

If we take Quine's topiary metaphor seriously, then cognitive science becomes the study of the forces that shape language and intelligence. Representations and language-like structures are a combination of external and internal pressures, whereby the formal characteristics of intelligent behavior are determined by the interplay between organisms in a social and physical environment. Of course such formal characteristics are constrained by the internal capacities of the biological mechanism; however, this constraint fits easily within the kinds of co-evolutionary framework discussed above. The topiary metaphor provides a way to understand how dynamicists can accommodate the obvious fecundity of representations in folk psychological contexts. Cognitive systems are responsive to representational systems in appropriate ways without themselves being representational systems.

The methodological insights of the dynamical movement mean that cognitive systems can be studied as self-modifying dynamical systems in response to standard patterns of intelligent behavior. The systematic properties of this intelligent behavior need not force us to look for some mechanism in the cognitive agent that generates patterns of intelligent behavior. We can recognize that cognitive systems exhibit the correct pattern of intelligent behavior in response to pressures from the

environment, without thereby assuming that an internal mechanism is responsible for this systematicity.

The kind of systematicity that had driven philosophers like Fodor to insist on the reality of inner representational structure, is to be found primarily in the social world and not in the brain. The development of the adult brain may be best thought of as a dynamical system following what Berg, Pollack and others have called a moving target strategy of development. But what about the adult brain itself? Clearly the adult brain has a set of capacities for the production of systematic behaviors of a characteristically intelligent kind. However, if we accept the moving target strategy, then any identification of representational structure in the brain will rest on a relatively arbitrary assignment of particular representational types to particular brain structures.

This general approach is consonant with the continued investigation of many of the neural phenomena currently labeled 'internal representations' by neuroscientists and cognitive scientists. Neuroscientists regularly use the term 'representation' for regularities in the brain that do not quite fit the standard criteria that characterizes uses of 'representation' in the philosophical literature (see for example Haugeland, 1991). Many of the useful and explanatory neural phenomena that neuroscientists call representations, for example, Eichenbaum and Cohen's hypothetical relational representations in the hippocampal system (1995) would not qualify as the kinds of thing that could be freely concatenated in a classical combinatorial syntax. Hence they do not play the kind of explanatory role that philosophers of mind expect of representations and are unlikely to qualify as the kinds of entities that feature in folk-psychological explanation. The approach outlined in this paper lifts most of the unwieldy conceptual burden from neuroscientific uses of the term 'representation' and suggests that we see neural 'representations' as adaptations to various patterns of meaningful activity.

7. Conclusion

This paper has presented a rough, initial hypothesis for the integration of work in dynamical systems theory with our ordinary folk-psychological patterns of explanation. However, rather than following Fodor's example, let's not assume that these patterns of intelligent behavior point to atomistic concepts and innate structures in the minds (or brains) of organisms. Instead, let's assume that these patterns are the manifestation of a relatively stable social and cultural landscape that the organism must negotiate in something like the way our limbs and muscles negotiate geographical landscapes. Happily, it's relatively easy for developing brains to manage, since the social and cultural environment for human beings has been shaped through natural history, by Baldwinian evolution, for the kinds of organisms we are. If we take the systematic properties of behavior as pointing to a shared social system rather than a set of rules for the brain to run its computations over internal representations, then a number of the basic problems that haunt the debate over Dynamical Systems theory simply dissolve.

8. Notes

¹Baldwin proposed that by temporarily adjusting behaviors during its lifespan, an animal could produce irreversible changes in the adaptive context of future generations. Though no new genetic change is immediately produced in the process, the change in conditions will alter which among the existing or subsequently modified genetic predispositions will be favored in the future (see Deacon, 1997, p. 322).

²For more on specific applications of DST to cognitive systems see Port and van Gelder (1996).

³Thanks to Yaneer Bar-Yam for helpful comments on these distinctions.

⁴More generally, systems may also be driven by flows of energy, this is not necessary for our discussion of the distinction between conservative and dissipative systems.

⁵Guinti (1995) shows that standard computational models, including the idealized Turing machine constitute a subset of possible dynamical systems. A Turing machine, for example can be identified with a mathematical dynamical system with discrete time: $\langle T, m, g^t \rangle$ where T is a set of non-negative integers, m is the set of triples $\langle \text{Tape content, head position, internal state} \rangle$, and the set of state transitions g^t is determined by the set of quadruples of the machine. (1995, 554–5).

⁶See also Clark, 1998, p. 117.

⁷See Smith (1998) for a valuable discussion of how the application of dynamical theories sheds light on the problem of approximate truth.

⁸We should note that such criticisms apply to any non-sentential kinematics of mental life.

⁹I think this is the point of many of Stich's papers in his *Deconstructing the Mind*.

¹⁰Of course, some philosophers of psychology with connectionist inclinations (most notably Horgan and Tienson, 89, 94) embrace dynamic and connectionist models of cognition while holding still that some kind of governance of cognition by representation. However, Garson has argued that Horgan and Tienson "risk instrumentalism" when it comes to syntax. I would argue that they should take the risk!

References

- Bechtel, W. (1998), 'Representations and cognitive explanations', *Cognitive Science* 22, pp. 295–318.
- Beer, R.D. (1995a), 'Computational and Dynamical Languages for Autonomous Agents', in R. Port and T. Van Gelder, eds., *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge, MA: MIT Press, pp. 121–147.
- Beer, R.D. (1995b), 'A dynamical systems perspective on agent-environment interaction', *Artificial Intelligence* 72, pp. 173–215.
- Berg, G. (1992), 'A Connectionist Parser with Recursive Sentence Structure and Lexical Disambiguation', *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI-92)*, 32–37.
- Brooks, R. (1991), 'Intelligence without reason' in *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, New York: Morgan Kaufmann.
- Chemero, A. (2000), 'Anti-representationalism and the Dynamical Stance', *Philosophy of Science*.
- Chomsky, N. (1959), A review of B.F. Skinner's *Verbal Behavior*, *Language* 35, pp. 26–58.
- Clark, A. (1997), *Being There*, Cambridge, MA, MIT Press.
- Cohen, N.J. and Eichenbaum, H. (1995), *Memory, Amnesia, and the Hippocampal System*, Cambridge, MA: MIT Press.
- Davidson, D. (1999), 'The Emergence of Thought', *Erkenntnis*.
- Deacon, T. (1998), *The Symbolic Species*, New York: Norton.
- Dennett, D. (1977), A cure for the common code. *Mind*. Reprinted in *Brainstorms* (MIT Press, 1978).
- Dennett, D. (1988), 'Précis of *The Intentional Stance*', *Brain and Behavioral Sciences* 11, pp. 495–546

- Edelman, G.M. (1987), *Neural Darwinism: The theory of neuronal group selection*, New York, NY: Basic Books.
- Eliasmith, C. (1997), 'The third contender: a critical examination of the dynamicist theory of cognition', *Journal of Philosophical Psychology* 9(4): 441–463.
- Fodor J. (1987), *Psychosemantics*, Cambridge, MA: MIT Press.
- Fodor, J. and Z. Pylyshyn (1988), 'Connectionism and cognitive architecture: A critical analysis', *Cognition* 28: 3–71.
- Fodor, J. and B. McLaughlin (1990), 'Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work', *Cognition* 35: 183–204.
- Freeman, W. (1988), 'Why neural networks don't yet fly: Inquiry into the neurodynamics of biological intelligence', *Proceedings of the IEEE international conference on neural networks* 2, 1–7.
- Freeman, W. (1990), 'Searching for signal and noise in the chaos of brain waves', in S. Krasner, ed., *The ubiquity of chaos* (pp. 47–55). Washington, DC: American Association for the Advancement of Science.
- Freeman, W. (1991), 'The Physiology of Perception', *Scientific American*, February, 35–41.
- Freeman, W.J. and Baird, B. (1987), 'Relation of Olfactory EEG to Behavior: Spatial Analysis', *Behavioral Neuroscience* 101, 393–408.
- Garfield, J. (1988), *Modularity in Knowledge Representation and Natural Language Understanding*, Cambridge, MA: MIT Press.
- Gargan, J. (1997), 'Syntax in a Dynamic Brain', *Synthese*, 110, 343–355.
- Horgan, T.E. and Tienson, J. (1993), 'Levels of Description in Nonclassical Cognitive Science', *Philosophy* 34, *Royal Institute of Philosophy Supplement*, 159–188.
- Horgan, T.E. and Tienson, J. (1996), *Connectionism and the Philosophy of Psychology*, Cambridge, MA: MIT Press.
- Kelso, J. A. S. (1995), *Dynamic Patterns: The Self-Organization of Brain and Behavior*, Cambridge, MA: MIT Press.
- Kelso, J.A.S., DelColle, J. and Schpner, G. (1990), 'Action-perception as a pattern formation process', in M. Jeannerod, ed., *Attention and Performance XIII*, Hillsdale, NJ: Lawrence Erlbaum Associates, 139–169.
- Newell, A. (1980), 'Physical Symbol Systems', *Cognitive Science* 4, 135–183.
- Pollack, J.B. (1991), 'The Induction of Dynamical Recognizers', *Machine Learning* 7: 227–252.
- Port, R. and van Gelder, T.J. (1995), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge, MA: MIT Press.
- Putnam, H. (1988), *Representation and Reality*, Cambridge, MA: MIT Press.
- Quine, W.V.O. (1960), *Word and Object*, Cambridge, MA: MIT Press.
- Skarda, C. and Freeman W.J., (1987), 'How brains make chaos in order to make sense of the world', *Behavioral and Brain Sciences* 10, 161–195.
- Smith, P. (1998), 'Approximate Truth and Dynamical Theories', *British Journal for the Philosophy of Science* 49, 253–277.
- Stich, S. (1996), *Deconstructing the Mind*, Oxford: Oxford University Press.
- Symons, J. (forthcoming), 'Co-evolutionary Models of the Emergence of Language'.
- Thagard, P. (1997) *Mind: An Introduction to Cognitive Science*, Cambridge, MA: MIT Press.
- Thelen, E. and Smith, L.B. (1994), *A Dynamic Systems Approach to the Development of Cognition and Action*, Cambridge, MA: MIT Press.
- Van Gelder, T.J. (1995), 'What might cognition be, if not computation?' *Journal of Philosophy* 91, 345–381.
- Van Gelder, T.J. and Port, R. (1995), 'It's About Time: An Overview of the Dynamical Approach to Cognition', in R. Port and T. van Gelder, eds., *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge MA: MIT Press.
- Van Gelder, T. (1998), 'The Dynamical Hypothesis in Cognitive Science', *Behavioral and Brain Sciences* 21, 1–14.