

1-1-2011

Anti-representationalism: Not a Well-founded Theory of Cognition

Michael Kirchhoff

University of Wollongong, kirchhof@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/artspapers>



Part of the [Arts and Humanities Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Kirchhoff, Michael, Anti-representationalism: Not a Well-founded Theory of Cognition 2011, 1-34.
<https://ro.uow.edu.au/artspapers/987>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Anti-representationalism: Not a Well-founded Theory of Cognition

By Michael David Kirchhoff,
The University of Wollongong.

Abstract: This article argues for the conclusion that anti-representationalism in the cognitive sciences is not a well-founded theory of cognition. This conclusion is supported by the observation that the link between the sceptical *demonstrations* and the anti-representational *conclusion* is too weak for the demonstrations to justify anti-representationalism *in general*. Rather than denying the need for internal representation, this article aim to establish that representational explanation - reconstructed within a dynamical agent-environment characterization - serves a necessary epistemic and ontological aim: It enables us to demarcate activities that presuppose *intentionality* and *behavioral autonomy* from activities that are merely reactive and situation-determined.

1. Introduction - the radical embodied cognition thesis

During a decade or so cognitive science has seen the rise of an *anti-representational* alternative attempting at unseating the concept of *internal representation* from its position as a core theoretical construct of scientific cognitive research. This anti-representational strand is historically rooted in continental phenomenology, especially the works of Heidegger and Merleau-Ponty. It is a direct attack on a formerly uncontested assumption: that internal representation distinguishes mind from the rest of the physical world. What is in dispute here is not that internal brain processes are important to cognition; it is the presence within us of *identifiable, well-individuated* vehicles of content that is being denied. Philosopher of cognitive science Andy Clark has summarized the anti-representational scepticism under the phrase, *The Radical Embodied Cognition Thesis*:

Structured, symbolic, representational, and computational views of cognition are mistaken. Embodied cognition is best studied using *noncomputational* and *nonrepresentational* ideas and explanatory schemes, and especially the tools of dynamic systems theory. (2001, p. 128; italics added).

There are at least two important arguments against representation in the radical

embodied cognition thesis. The first of these arguments, the threat from *nontrivial causal spread*, occurs whenever the material vehicles of cognitive architecture are *causally spread* beyond the brain and *nontrivially* involved in the completion of cognitive tasks. The second of these arguments, the threat from *continuous reciprocal causation*, occurs whenever the causal contributions made by components of a system partially determines and is partially determined by causal contributions of other systemic components, thereby making it impossible to assign a *specific* subtask to an *identifiable* subsystem within a larger system. None of these arguments are derived from arm-chair reflections, but are supported by empirical evidence from the major disciplines and fields of research involved in the investigation of cognition: neuroscience (e.g., Freeman, 1999), ecological psychology (Gibson, 1979), developmental psychology (e.g., Thelen & Smith, 1994), situated robotics and autonomous agent theory (e.g., Brooks, 1990, 1991, 1999; Hendriks-Jansen, 1996), philosophy of mind (e.g., Keijzer, 1998; van Gelder, 1995, 1998; Wheeler, 1994, 1996, 2001, 2005ab), the dynamical approach to cognition (e.g., Haselager et al, 2003ab, Kelso, 2003; Thelen, 2003), and computational neuroethology (e.g., Beer, 1995; Chiel & Beer, 1997; Webb, 1994, 1996).

1.2. Paper thesis and its underwriting argument

The anti-representational challenge to cognitive science is apt and of immense importance to the development of future models of cognition. But, as I shall show here, it is by no means free of trouble – hence the reason for this article. In the article I will show that the link between the empirical *demonstrations* and the sceptical *conclusion* is too weak for the demonstrations to justify the anti-representational conclusion. The argument put forth in support of this thesis turns on the hypothesis that anti-representationalism is not a well-founded theory of cognition. The logical structure of this argument is:

- i. In order for a theory, (**T**), to count as a well-founded anti-representational theory of cognition, (**T**) must involve *no* use of representation-laden

- concepts, and (*T*) must have conceptually *sufficient* means to explain cognitive phenomena *in toto*.
- ii. All of the anti-representational demonstrations are merely *reactive* in character, in the sense that all of these demonstrations only pertain to systems that are *uninterestingly non-cognitive*.
 - iii. Cognitive properties are emergent properties, but emergent properties – together with the tools of dynamical systems theory – are *not intrinsically* nonrepresentational. *Therefore*:
 - iv. Anti-representationalism is *not* a well-founded theory of cognition.

Clark has made a similar claim, also targeting an apparent lack of fit between the scope of the case studies employed and the scope of the sceptical argument (see Clark, 1994, 1997). Although clearly motivated by the line of reasoning undertaken by Clark, the strategy put to use here is different. Where Clark points out that the kinds of problem-domains involved to justify the sceptical conclusion are insufficiently “representation-hungry” for them to do so (see Clark & Toribio, 1994, p. 418), the present paper goes on to argue an epistemological point - that anti-representationalism is *not a well-founded theory of cognition* - and an ontological point - that the anti-representational demonstrations only pertain to systems that are *uninterestingly non-cognitive*. It is easily documented that anti-representationalists do indeed make the transition - moving *from* empirical demonstrations *to* what is taken as a constitutive claim about the nonrepresentational character of cognition in general. Consider, e.g., the following claims by leading anti-representationalists:

Dynamics forms a powerful framework for developing models of cognition that sidestep representation altogether. The assumption that cognition must involve representation is based in part on inability to imagine how any nonrepresentational system could possibly exhibit cognitive performances. *Within the dynamical approach, such systems can not only be imagined, they can be modelled and constructed.* (van Gelder, 1998, p. 622; italics added).

We conclude here that as all mental activity is emergent, situated, historical, and embodied, there is *in principle no difference* between the processes engendering walking, reaching (...) and those resulting in mathematics and poetry. (Thelen & Smith, 1994, p. xxiii; notes omitted; italics added).

(...) there is good evidence that when evolution is given the right evolutionary building blocks (...), the result will often be control systems that feature high degrees of continuous reciprocal causation. Where that is so, those control systems will be stubbornly resistant to (...) representational analysis. (Wheeler, 2005a, p. 229; notes omitted).

The strategy of the rest of this paper is accordingly. I will expound my argument in four steps. First, I will unpack the notion of representation. Second, I will deal exclusively with the argument from nontrivial causal spread by focusing on programmatic presuppositions and empirical cases. Third, I will discuss the threat from continuous reciprocal causation. Finally, I will advance and defend the thesis of this article.

2. *Defining representation*

Any attempt at defining the concept of internal representation needs to take into account an important change of inquiry on representation, which is that the debate has changed from being about the *content* of representation to now being a debate about *the role that representations are supposed to play in behavioral adaptivity and intelligence*. In the light of recent anti-representational scepticism, it is this latter kind of inquiry that will be my focus here. There exist two types of representational accounts in cognitive science and philosophy of mind: accounts based on the criteria of *decoupling* - i.e., roughly, the separability of representations from their stimulus - and *teleological* definitions - i.e., roughly, the evolutionary, functional role of a representation in adaptive behavior.

2.1. *Decoupling based account of representation*

If one holds an account of internal representation based on the criteria of decoupling, then a system will be a representational using system if, and only if, the three following conditions are upheld:

- The system must exhibit the presence of inner states whose functional role is to “*stand in*” for environmental features (Haugeland, 1991, p. 62; Smith, 1996, p. 220).
- The system must allow for “*precise identification*” of inner states or processes

- with representational roles (Clark, 1997a, p. 147; Wheeler, 2005b, p. 224).
- The inner states so identified must “*enhance*” fluent real-time action (Clark & Grush, 1999, p. 7; Haugeland, 1991, p. 62).

2.2. Degrees of decoupleability

For those committed to such decoupling requirements for representation, the next (contemporary) question of scientific interest turns on the *degree* of decoupleability necessary for an inner state to qualify as an inner *representational state*. Some theorists hold the view that for (**R**) to count as a representation of its target (**T**), (**R**) must be capable of standing-in for *temporally and spatially* absent states of affairs in the world (see e.g., Haugeland, 1991, p. 62; Newell, 1980, p. 142-47; Smith, 1996, p. 220; and Sterelny, 1990, p. 20-21). On this version of decoupleability, (**R**) and (**T**) must *not be in constant causal contact*, (**T**) must, in some cases at least, be *fully* absent, and (**R**) must be *adaptive*. This strong version of decoupleability, some theorists do not accept. In fact, a second account of the decoupleability definition of representation denies that decoupleable inner states, in the strong sense, are necessary for representation, since, so the argument is formulated, it rules out cases where inner states function only so as to control immediate environmental interaction (see e.g., Ballard, 1991; Clark & Grush, 1999; and Dorffner, 1999). On the weak (or moderate) view, it is sufficient that (**R**) only stand-in for *temporally* absent states of affairs. Hence, here (**R**) and (**T**) must *not be in constant causal contact*, (**T**) must be *temporally (not spatially) absent*, and (**R**) must be *adaptive*.

2.3. Two types of decoupleability accounts

Holding a decoupling based account of representation has *metaphysical* consequences that go along with the conceptual ones just sketched. Metaphysical issues come forth when considering the *relation* between internal representation and reality (world). On the decoupling side, there exist two versions of how internal states are supposed to be related to worldly states of affairs. On the one hand there are those advocating a *referential* relation between inner state and

reality and there are those promoting an *interactional* relation on the other. The primary assumption of the *referential* view is that there is a *mind-independent* reality which must be encoded onto a model in the mind of an agent in order for the agent to act intelligently (see e.g., Fodor & Pylyshyn, 1988; Newell, 1980; Pinker, 1997; Sterelny, 1990; and others). Now, if one's theory of representation is *interactional* in character, then for (\mathbf{R}) to be a representation, the *relation* between (\mathbf{R}) and its target property (\mathbf{T}) is believed to be an *egocentric, non-objective* and *action-oriented* relation (occasionally this relation is also called 'internal relation', picked up from classical ontological terminology by e.g., Bickhard (2004)). Consequently, this particular *kind* of representational relation differs substantially from its referential alternative, because interactional representations do not *refer* to anything in reality; rather, they *constitute* reality for the organism (see e.g., Ballard, 1991; Chemero, 2000; Clark, 1997ab; Costa & Rocha, 2005; Damasio, 1994, 2000; Dorffner, 1999; Kirsh, 1995, 1996; Peschl & Riegler, 1999; Smith, 1996; and others).

2.4. Teleological based account of representation

On the opposite side of the representational landscape, there are those who claim that insisting on decoupling as a necessary condition for representation is *controversial*. On this view (\mathbf{R}) will count as a representation of its target (\mathbf{T}), when (\mathbf{R}) and (\mathbf{T}) are in *constant causal contact* with one another, when (\mathbf{T}) is *present*, and when (\mathbf{R}) is adaptive *for the system itself*. Decoupling, therefore, is conceived as only sufficient for representation and not as a necessary criterion. This sort of view is often called a *teleological* account of representation (see e.g., Bickhard & Terveen, 1995; Bickhard, 2003, 2004; Chemero, 2000; Christensen & Hooker, 2004; Millikan, 2004). On a teleological definition, what constitutes that (\mathbf{R}) of a system (\mathbf{S}) will count as a representation *for* (\mathbf{S}) has *nothing* to do with what (\mathbf{R}) is or is not related to – or decoupled from. In fact, (\mathbf{R}) will count as a representation *if, and only if*, (\mathbf{R}) stands between a representation *producer* and a representation *consumer*; (\mathbf{R}) is functionally adaptive with respect to a consumer system; and that (\mathbf{R}) is part of a larger representational system (see Chemero, 2000, p. 627;

Millikan, 2004, p. 75).

3. *Extended cognition & nontrivial causal spread*

The phenomenon of *nontrivial causal spread* surfaces as a kind of background worry in a myriad of scientific fields, all of which revolve around the extended cognition thesis: That cognitive processes are *not* located *inside* an agent *exclusively* and that cognition is *co-constituted* by such extended physical vehicles. Because of this, the locational claim and constitution claim of extended cognition results in anti-representationalism whenever the physical vehicles of cognition are *causally spread* beyond the brain and *at the root of* a particular cognitive task. In principal, the sceptical argument advanced from nontrivial causal spread is a direct attack on the *responsibility requirement* of inner representational accounts: In order for (**R**) to justify as a representation - on either the decoupling or the teleological account - (**R**) must be causally *responsible* for the behavior in question. According to Wheeler & Clark (1999) and Wheeler (2001), nontrivial causal spread may be defined as follows:

Causal spread obtains when some phenomenon of interest turns out to depend, in *unexpected* ways, upon causal factors *external* to the system *previously/intuitively* thought *responsible*.” (1999, p. 105; italics added). “In the context of cognitive science, non-trivial causal spread exists when additional factors reveal themselves to be the *unexpected root* of the *very adaptive flexibility and richness* that is *normally* attributed to representation-based control. (2001, pp. 217-218; italics added).

My ambition is not simply to illuminate the insights gathered in Wheeler & Clark (1999) and in Wheeler (2001, 2005ab). I aim to move beyond these authors, since, I submit, they do not provide a sufficiently nuanced description of the underlying intuitions plus arguments, which either motivates or underwrites anti-representational scepticism ignited by cases of nontrivial causal spread. When reading these authors, one is often left with the impression that only one *kind* of phenomenon signifies nontrivial causal spread – namely, cases where extra-neural processes take on the role previously thought to be restricted to brain states/processes alone. Clearly this is *one* kind of phenomenon, where nontrivial

causal spread is implicated. It's just not the *only* one. To make sense of the different assumptions and arguments involved in nontrivial causal spread style theorizing, I propose setting up a taxonomy. By grouping this anti-representational phenomenon into different classes, my hope is to shed some light on the nature of nontrivial causal spread.

3.1. Taxonomy for nontrivial causal spread

There are two different kinds of arguments lending themselves *as* reasons for anti-representational scepticism, and both of these are embedded in the phenomenon of nontrivial causal spread. One *kind* of argument is rooted in the “*principle of parity*” (e.g., Thelen & Smith, 1994; Wheeler, 2001, 2005b), whereas another argument turns on what I will call the “*coupling as constitution*” argument (e.g., Gibson, 1979; Keijzer, 1998; Thelen, 2003; Webb, 1996).

First, if (\mathbf{X}) – an extra-neural factor/process – and (\mathbf{Y}) – a neural factor/process – are so coordinated that they *together* constitute some psychological phenomenon (\mathbf{P}), and the argument invoked to make this assertion is underpinned by the “principle of parity”, then one will hold the view that there is *no principled difference* between (\mathbf{X}) and (\mathbf{Y}) in their contribution to (\mathbf{P}). On this view, nontrivial causal spread may be understood to express the following claim: If it is *equally credible* to assign the same functional role to (\mathbf{X}), as we normally or intuitively do to (\mathbf{Y}), then (\mathbf{X}) is a part of the cognitive process constituting (\mathbf{P}). Consider the following remark by Thelen & Smith:

There is, (...), *no essence* of locomotion either in the motor cortex or in the spinal cord. Indeed, it would be *equally credible* to assign the essence of walking to the treadmill than to a neural structure, because it is the action of the treadmill that elicits the most locomotor-like behavior. (1994, p. 17; note omitted; second italics added).

The parity principle is particularly clear in this remark: If a part of world functions as a process, were it performed by the brain, we would not have trouble recognizing it as a cognitive process, then that part - in this case the treadmill - is part of the cognitive process. From this empirical case, Thelen & Smith infer anti-representationalism, evident in the following quotation:

How do minds change? Where does new knowledge, (...), new behavior come from? How does the organism continually adapt and create new solutions to new problems? The answer we present (...) makes *no use of representations or representation-like processes*. (1994, p. 42; notes omitted; italics added).

Second, if (\mathbf{X}) and (\mathbf{Y}) are so coordinated that they *together* constitute (\mathbf{P}), and that the argument for this kind of nontrivial causal spread turns on what I have termed the “coupling as constitution” argument, then one will tend to derive anti-representational scepticism from the view that (\mathbf{X}) and (\mathbf{Y}) *jointly govern* the processes leading to (\mathbf{P}). That is, one will entertain the view that (\mathbf{X}) and (\mathbf{Y}) make up a *coupled system*. More specifically, if (\mathbf{X}) and (\mathbf{Y}) form a causally coupled system, then neglecting to take (\mathbf{X}) into account when explaining (\mathbf{P}) necessarily leads to a failure in recognizing a vital aspect underpinning the generation of (\mathbf{P}). Hence, (\mathbf{X}) and (\mathbf{Y}) are *causally* and *constitutively* coupled to each other. Contrary to the scepticism motivated by the principle of parity, the coupling as constitution argument does not turn sceptical from a position of no-principled-difference. Instead it turns on the co-constituency of each causal component involved. Among those explicitly giving voice to this kind of nontrivial causal spread are Keijzer, Thelen and Webb:

Yet again it appears that it is the interaction of the robot’s *uncomplicated* mechanisms with particular *sound fields* that produces this interesting – and useful – behavior.” (Webb, 1996, p. 67; italics added). “Behavior is an emergent pattern of *cooperating* components, *all* of which count and *none* of which are privileged. (Thelen, 2003, p. 20; italics added).

(...) the notion of *representation is to be dispensed with*. Instead, behavior is to be explained as the *intricate* interaction between an embodied organism and the specific make up of an environment. (Keijzer, 1998, p. 269; italics added).

Keeping this distinction between nontrivial causal spread motivated by the principle of parity and coupling as constitution in mind, let us turn to look at a couple of empirical case-studies.

3.3. Infant stepping behavior

Here is a compelling case of nontrivial causal spread motivated by the principle of parity, from research on the development of coordinated stepping behavior in

human infants. In spite of being motorically immature, infants will, when held in an upright position, *precociously* lift and lower their legs in a coordinated and step-like fashion (Thelen & Smith, 1994, pp. 10, 89). But, at approximately 2 months of age, this alternating pattern of stepping behavior disappears from the infants' repertoire of motor skills. First at about 8 to 10 months, when the ratio between muscle- mass and non-muscle tissue (fat) has settled into equilibrium, do the infants regain their *prior* ability to perform such stepping movements. The final milestone is reached at around 12 months of age, where the first independent steps become possible (Thelen & Smith, 1994, p. 10). In developmental psychology, the question is how this *universal* developmental pattern might be *reliably* explained?

In a series of empirical experiments, Thelen & Smith found that during the non-stepping window between 2 to 8 months, 7 month old infants *will* perform stepping movements, *if* they are supported by a *slow-moving, motorized treadmill* (1994, pp. 12, 94). The hypothesis derived from this observation is that the mechanical action of the treadmill elicits a shift between non-stepping and stepping behavior (Thelen & Smith, 1994, p. 96). On their account, the treadmill may do so, because it *substitutes* for changes in leg dynamics that occur naturally in adult locomotion. Right after the first foot touches down on the treadmill, the infant's center of gravity is shifted to its stance leg, and the other leg is pulled or stretched backwards. This *stretch* is important *informationally*. When the trailing leg is fully stretched, this will provide the CNS with *proprioceptive feedback* triggering the initiation of a forward swing (Thelen & Smith, 1994, p. 96). So, when supported on a treadmill, the treadmill will elicit a pattern of self-organizing steps "(...) to complete the loop and allow the pattern [of coordinated stepping behavior], normally cryptic, to become manifest." (Thelen & Smith, 1994, p. 97; note omitted; square bracketed quotation added).

This conclusion makes fully explicit why Thelen & Smith's experiment on treadmill stepping is a demonstration of nontrivial causal spread motivated by the principle of parity. It demonstrates that there is *no principled difference* between neural and non-neural factors in their contribution to stepping behavior. If it is

equally credible to assign the same functional role to the treadmill, as we normally or intuitively do to neuronal processes, then the treadmill is *part and parcel* of the cognitive process constituting coordinate stepping behavior. Moreover, it is a clear cut example of Wheeler's emphasis on *scientific discovery* – contrary to an entrenched understanding, motor development, or stepping behavior, is not governed by neural mechanisms only, but is *causally spread* to factors *spatially distanced* from prior neuro-centric intuitions (2005b: 218).

3.4. Cricket phonotaxis

Another compelling case of nontrivial causal spread, although this time underpinned by the coupling as constitution argument, comes from computational neuroethology – especially Barbara Webb's (1994, 1996) work on cricket phonotaxis. Male crickets produce, by rubbing one wing against the other, a *species-specific* song in order to attract a female cricket. The ability of female crickets to identify, locate and move towards a specific song, along with all the sensorimotor abilities such an activity entails, is known as *phonotaxis* (Webb, 1996, p. 63). Anatomic, neural and environmental details are required for us to appreciate this unique skill:

A female cricket has two ears, one on each of its two front legs. The cricket's body channels sound through a tracheal tube of fixed length which connects the two ears to one another and to further openings, called spiracles, on the cricket's body (Webb, 1996, p. 63). Thus, sound arrives at the ear-drums both from an *external source* (from the male's signal) and from an *internal source* (via the tracheal tube). Because of the difference of *route*, a female cricket informationally picks-up a male crickets sound accordingly: At the ear closer to the sound source, the external sound travels less distance than the internal sound arriving via the trachea, whereas, on the side further away from the sound source, the two sounds travel the same distance (Webb, 1996, p. 66). Because of the difference in *distance*, the external and internal arriving sounds are *out-of-phase* at the eardrum on the side closer to the sound source, whereas they are *in-phase* at the eardrum located further away. What this means is that the *amplitude* of

eardrum vibration is *greater* at the eardrum *closer* to the sound source (Webb, 1996, p. 66). According to Webb, the significance of this is immense, because it is the higher pitch in eardrum amplitude which elicits the direction of travel chosen by the female cricket (1996, p. 63). The direction of travel is achieved via the *firing rate* of a pair of *inter-neurons* in the crickets' CNS – one connected to the right ear and the other to the left ear (Webb, 1996, p. 64). Which inter-neuron fires first is dependent on the strength in amplitude at the eardrums. So, the scientific conjecture is that the female crickets CNS triggers a turning toward the side on which the sound source is closer, because at the eardrum closer to the sound source the pitch in amplitude will be higher than at the eardrum located further away (Webb, 1996, p. 64). Here the specific *temporal pattern* of the male's signal becomes important, since if the firing rate of each inter-neuron is sensitive to a certain pitch in amplitude at the eardrum, then the *frequency* of a male crickets sound burst must be important. In fact, what Webb found was that the frequency in sound bursts – known as *syllables* – is the most essential feature in eliciting a female response (1996, p. 63). Neither must the period between syllables be too long nor too short: too long period's lead to an informational infrequency from poverty in environmental stimuli; and too short period's lead to an uncertainty as to whether which one of the inter-neurons fired first.

From this example, Webb goes on to derive the insight that apparently complex behavior of an agent does not imply complexity of the underlying mechanism (1996, p. 66). Rather, adaptation in female crickets arises from interactions *nontrivially causally spread* across a crickets CNS – the inter-neurons – its body – the fixed length tracheal tube and spiracles – and its environment – the syllables of the male cricket's species-specific song. Furthermore, if one fails to recognize any of these constituents, thus the argument goes, it's the equivalent of failing to recognize a vital part of natural phonotaxis.

4. Dynamics & continuous reciprocal causation

While the argument from nontrivial causal spread downplays the kind of contribution that neural structures/processes make to the generation of adaptive

intelligent behavior, this second sceptical argument advances an empirical (or scientific) hypothesis about cognition. It suggests that cognitive phenomena emerge from *continuous, interdependent and circular causation* at the level of systemic components. The attack is directed at the *locatability requirement* of representational accounts: In order for a system (\mathcal{S}) to justify as a representational system, (\mathcal{S}) must contain *identifiable* sub- systems that perform *specifiable* subtasks. Such is the threat from the specific mode of causation identified as *continuous reciprocal causation* (the term is due to Clark, 1997a).

4.1. *Exposing the argument of continuous reciprocal causation*

I will begin by exposing the argument in its logical form. Having achieved this, the next job will be to present some of the pivotal presuppositions made in the argument. Finally the argument, and the assumptions singled out, will be elaborated. The argument immanent in the threat from continuous reciprocal causation is as follows (e.g., van Gelder, 1995, p. 351; Wheeler, 2005b, pp. 223-29):

- i. Functional homuncularity is necessary for representation.
- ii. Modularity is necessary for functional homuncularity.
- iii. Many physical systems underlying online (adaptive) intelligence are not modular. *Therefore:*
- iv. Many physical systems underlying online intelligence are not homuncular. *Therefore:*
- v. Many physical systems underlying online intelligence are not representational.

4.2. *Homuncularity, modularity & representation*

To cut nature at one of its joints – by taking psychological kinds to be representational – while at the same time advocating the view that systemic behavior is causally-functionally explicable in terms of the properties of and interactions among the properties of that system is the characteristics of *orthodox* cognitive science. It involves appeal to functional homuncularity, modularity and

representation (see e.g., Fodor, 1983, pp. 38-66; Sterelny, 1990, p. 13). Wheeler is particularly clear on this matter. Functional homuncularity is a subproperty of modularity and functional homuncularity is necessary and sufficient for internal representation (2005a, pp. 252-60). Thus, if continuous reciprocal causation prevents homuncular analysis, and homuncularity is a subproperty of modular systems, then to defy modular analysis is also to resist representational explanation (see Beer, 1995, p. 208). So is the *explicit* sceptical presupposition. Clarification is required:

Functional homuncularity denotes a subset of modular systems. Modularity of a system property – e.g., visual perception – presupposes that the system property, (*Sp*), consists of scientifically *identifiable* subsystems, (*Sp1, Sp2, ..., Spn*), each of which performs a clear, well-defined function – e.g., face-recognition – and that the overall function of (*Sp*) is dependent on the mode of organization among (*Sp1, Sp2, ..., Spn*). So, to explain the functioning of a complex whole by individuating its parts and detailing their overall organization is to engage in modular explanation (see e.g., Fodor, 1983, p. 1). Now, if a modular system, (*Sp*), is eligible for homuncular explanation, then (*Sp*) can be *hierarchically* decomposed into a set of organized *communicating* subsystems (*Sp1, Sp2, ..., Spn*), where (*Sp1, Sp2, ..., Spn*) each contributes to the collective achievement of the overall function of (*Sp*). So functional homuncularity cuts nature slightly different than does modularity, because the former says something very specific about the latter – namely, that the subsystems, (*Sp1, Sp2, ..., Spn*), interact by *passing information to each other*. The anti- representational case is directly derivable: *If continuous reciprocal causation undermines modularity, and a necessary requirement for representation entails scientific identification of task- specific subsystems, then the presence of continuous reciprocal causation undermines the locatability requirement necessary for homuncular interpretation.*

4.3. *Emergence, dynamic systems & continuous reciprocal causation*

A shared implicit assumption of the sceptical argument is (very roughly) the empirical hypothesis that certain target properties – including cognitive and

psychological ones – are *ontologically* and *epistemologically* best conceived of as “*emergent* [properties] of nonlinear *dynamical systems* (...)” (van Gelder, 1998, p. 616; square bracket quotation added; note omitted). Strictly speaking, what the assumption implies is that “cognitive properties” are a subset of “dynamical properties”, and that “dynamical properties” are “emergent properties”. In fact, if a cognitive property is an emergent property, then functional homuncularity and modularity are problematic.

A defining feature of emergence, as the concept is put to use in the anti-representational literature, is that emergence of a systemic property is a *failure* of modularity of a systemic property. This tells us something important about being emergent. First, emergent properties are systemic properties – i.e. that, a property is emergent if and only if the system possesses it but *no single component* of the system possesses it (see e.g., Stephan, 2006, p. 487). Hence, an emergent systemic property is a *collective property*; it is not a property of any single part of a system. Second, an emergent property depends on continuous reciprocal causation among its subsystemic components (van Gelder, 1995, p. 353). So, if property (*Sp*) is emergent, then (*Sp*) results from circular, interdependent change among its subsystemic parts (*Sp1, Sp2,..., Spn*). From these two features of emergence, some – for instance, Wheeler (2005b, p. 225) – go on to claim that if (*Sp*) is an emergent property, then this prevents explanation of (*Sp*) in terms of its properties (*Sp1, Sp2,..., Spn*) and their mode of organization. The implicit assumption of the argument is that cognitive properties are emergent properties, and whenever a property of a system is emergent, it will defy functional homuncularity, modularity, and in the end representation.

4.4. *Continuous reciprocal causation & the dynamical hypothesis*

Van Gelder has presented just such a sceptical argument. First, ontologically, he claims that “cognitive systems may in fact be dynamical systems, and cognition the behavior of some (noncomputational) dynamical system.” (1995, p. 358). Second, epistemologically, van Gelder claims that the correct explanatory tools of cognitive science are those of dynamical systems theory (DST). This is what

makes up the *dynamical hypothesis* (DH) in cognitive science. Carrying the evidential load is a central example of a dynamical system – the Watt centrifugal governor. The claim he wishes to make is that the Watt governor *can* and *should* be considered as a prototypical model for cognition (1995, p. 347).

The governor was designed by Scottish engineer James Watt circa 1788 as a solution to the problem of keeping constant the speed of a flywheel that drives industrial machinery, while itself being driven by a steam engine. Since the speed of the flywheel is affected both by steam pressure, and by current workload placed on the engine (no. of machines being driven, etc.), the speed of the flywheel tend to fluctuate. In order to control the flywheel, one must control the amount of steam entering the pistons from the engine via a throttle-valve. A centrifugal governor is a device which can automatically adjust the throttle value so as to maintain constant speed of the flywheel in spite of continuous fluctuations in steam pressure and work load. Watt's solution was as follows:

It consisted of a vertical spindle geared into the main flywheel so that it rotated at a speed directly dependent upon that of the flywheel itself. Attached to the spindle by hinges were two arms, and on the end of each arm was a metal ball. As the spindle turned, centrifugal force drove the two balls outwards and upwards. By a clever arrangement, this arm motion was linked directly to the throttle valve. The result was that as the speed of the main flywheel increased, the arms raised, closing the valve and restricting the flow of steam; as the speed decreased, the arms fell, opening the valve and allowing more steam to flow. The result was that the engine adopted a constant speed, maintained with extraordinary swiftness and smoothness in the presence of large fluctuations in pressure and load. (Van Gelder, 1995, p. 349).

First, ontologically, van Gelder claims that the operation of Watt's design solution is *nonrepresentational in nature* (1995, p. 351). This is the “nature hypothesis” of the DH: Cognitive agents are, like the governor, dynamical systems, and dynamical systems are nonrepresentational in character. He advances four different arguments in support of this assertion; only the fourth and most interesting will concern us here. Representation presupposes functional homuncularity. If a system is homuncular, then it's necessarily modular. The Watt governor, however, does not “exhibit this cluster of properties as a whole,

nor any of them individually.” (Van Gelder, 1995, p. 351). So, the governor cannot be representational. Against modularity van Gelder asserts that there is no possibility of *nonarbitrarily* cutting the governor’s systemic property, (*Sp*), into scientifically identifiable subsystems (*Sp1, Sp2,..., Spn*), because (*Sp*) is an emergent property of continuous, simultaneous and interdependent interactions among its microstructural parts:

(...) arm angle and engine speed are at all times both determined by, and determining, each other’s behavior. (Van Gelder 1995: 351; note omitted).

Since functional homuncularity is a subproperty of modular systems, showing that the Watt governor prevents modular analysis is sufficient in order to underpin the claim that the governor is nonrepresentational in character (van Gelder, 1995, p. 354). Therefore, to deny modularity is the equivalent of denying homuncularity, which, in turn, prevents representational analysis:

(...) when we fully understand the relationship between engine speed and arm angle, we see that the notion of representation is just the wrong sort of conceptual tool to apply. (Van Gelder 1995: 353; note omitted).

These properties – representation, computation, sequential and cyclic operation, and homuncularity – form a mutually interdependent cluster; a device with any one of them will standardly possess others. Now, the Watt centrifugal governor does not exhibit this cluster of properties as a whole, not any one of them individually. (van Gelder, 1995, p. 351).

Second, epistemologically, van Gelder claims that only the conceptual tools of DST can explain the circular causation between arm angle and engine speed. This is the “knowledge hypothesis” of the DH: Cognitive science can and should take dynamical form, and dynamics forms a powerful framework for sidestepping representation altogether (1998, pp. 619-22). Formally, a *dynamical system* is any system for which we can provide a finite number of *variables* (a variable is some entity of a dynamic system subject to nonlinear change) and a set of *differential equations* (equations involving a function and one or more derivations) describing how the values of those variables change interdependently over time (see e.g., Wheeler, 1996, pp. 222-225). DST is the formal framework providing such explanatory tools for describing the evolution of dynamical systems over time.

The principles governing change in arm angle over time may be described by using the following differential equation (van Gelder, 1995, p. 356):

$$d^2\theta/dt^2 = (nw)^2 \cos \theta \sin \theta - g/l \sin \theta - r d\theta/dt$$

Where θ is the angle of the arms; n is a gearing constant; w is the speed of the engine; g is a constant for gravity; l is the length of the arm; and r is a friction constant. This equation describes the *nonlinear* change in arm angle as a function of the current arm angle θ , the way the arm angle is changing already (the derivative of θ with respect to time, $d\theta/dt$), and the current engine speed w . Of special interest to the knowledge hypothesis is that the equation shows that change in arm angle is an “emergent property” of *circular causation* among the systems ($d^2\theta/dt^2$) *subsystemic variables*. An important aspect of DST here is that it conceives of cognition – and dynamical systems in general – as the unfolding of complex structures over time (van Gelder, 1998, p. 621). For instance, the role of θ at time t^1 is determined (and helps determine) the roles of the other components (w and $d\theta/dt$) at t^1 , and may due to the inherent nonlinearity of the system contribute quite differently at t^2 in virtue of causally circular and self-organizing links to other components. DST, it seems, provides us with tools for understanding interactive, emergent phenomena – phenomena, so the argument is set up, too complex for functional homuncularity, modularity and representation. From this equation, van Gelder moves on to consider the steam engine itself as a dynamical system *causally coupled* with the governor (1995, p. 357):

$$d^2/dt^2 = F(w, \dots, t, \dots)$$

Where t is the current setting of the throttle valve, which depends directly on the arm angle, θ , of the governor. Just as w is a *parameter* – i.e., something outside of a system, but on which the system causally depends – of the governor system above, so θ is a parameter the engine system (van Gelder, 1995, p. 357). In this equation, the governor and the engine system form a self-organizing system of coupled *oscillators* – an oscillator is any system which executes periodic behavior (e.g., the flywheel, a neural rhythm, etc.) – in which both θ and w function as

variables. This mutual influence of two dynamical systems is called *coupling* – i.e., mutual direct interdependence among oscillators. In cases of extended cognition, this means that neither system (brain, body and world), considered separately, is adequate to explain some cognitive phenomenon. Rather, just like the governor and engine system make up a self-organizing system of coupled oscillators, so should one conceive of the relationship between brain, body and environment in terms of coupled, self-organizing oscillators. Strictly speaking, what this suggests, according to van Gelder, is that a framework which invokes homuncular decomposition of a system into a set of representational subsystems is fundamentally insufficient when having to account for the complex interplay of coupled, dynamical systems (1995, p. 369).

5. *Anti-representationalism: not a well-founded theory of cognition*

Based on the review of anti-representational arguments, in the previous sections, let us now turn to a critical discussion of the actual significance of these arguments. Let us begin with the observation that organismic (or animal) activities fall (roughly) along a continuum. At the one extreme are purely *reactive, situationally-determined* activities – e.g., tracking the ambient energy of the sun, treadmill elicited stepping behavior, phonotaxis, avoiding collisions, and so forth. At the other extreme are highly *cerebral, situation-independent* activities – e.g., counterfactual predication, abstract reasoning, concept learning, etc. In the middle of this continuum is a class of *real-time, representation-hungry* activities, where behavioral success is neither fully determined by ambient environmental stimuli nor fully disconnected from ongoing embodied, situated activity – e.g., adaptive, anticipatory behavior, foraging behavior in bees, responding to situational saliency, reasoning about the absent (e.g., grooming behavior of rhesus macaques), expert drivers, tennis players, carpenters, goal-directed movements, etc.

5.1. *The scope of the argument*

It is my conjecture that the demonstrations invoked to drive the sceptical

conclusion all fall within the class of purely reactive, situationally-determined activity. Where intelligent behavior is generated by reactive, situationally-determined activity, explanations that appeal to representations are, I believe, undermined. Either extraneural factors will be *at the causal root of* the observed behavior, thus favouring the conclusion of the threat from nontrivial causal spread. Or the behavior of such a reactive system will result from continuous reciprocal causation among its subsystemic components, making it scientifically hard to assign a specific task (perhaps representational task) to a specific subsystem. It is from this conjecture that I derive the thesis that the link between the empirical *demonstrations* and the sceptical *conclusion* is too weak for the demonstrations to justify the anti- representational conclusion. Logically the arguments has the following form:

- i. In order for a theory, (\mathbf{T}), to count as a well-founded anti-representational theory of cognition, (\mathbf{T}) must involve *no* use of representation-laden concepts, and (\mathbf{T}) must have conceptually *sufficient* means to explain cognitive phenomena *in toto*.
- ii. The anti-representational (empirical) demonstrations are merely reactive in character, in the sense that all of these demonstrations only pertain to systems that are uninterestingly non-cognitive.
- iii. Cognitive properties are emergent properties, but emergent properties – together with the tools of dynamical systems theory – are not intrinsically nonrepresentational. *Therefore:*
- iv. Anti-representationalism is *not* a well-founded theory of cognition.

This argument deals with the threats from the DH and continuous reciprocal causation as well as with extended cognition and nontrivial causal spread. The DH asserts that there is a large class of dynamical systems for which representational glosses have no explanatory utility, and that the best explanations of cognitive phenomena fall within this class. The same type of case is made by those seeking to marry the extended cognition thesis with nontrivial causal spread: If it is the case that extraneural processes are at the causal root of a

phenomenon, representational analysis will yield no extra value. Because of this, both sceptical arguments are best understood as involving an inference *from* a model (the Watt governor, treadmill stepping, phonotaxis...) *to* the more radical claim of anti- representationalism *in general*. Ruling in favour of parsimony, I will accept the premise that there is indeed a large class of systems for which representational explanation have no explanatory (scientific) utility. However, I will not accept that it follows from this premise that one must buy in to the much stronger conclusion: That the best explanations of cognitive phenomena fall within this class. My hypothesis may be schematized accordingly:

Here is a model of either some extended system or some nonlinear, dynamical system - for instance, the WG, infant treadmill stepping, etc. There are no representations in any of these models [*This is right*]. So, if cognition in general really works like these models do, then there are no representations in cognition either. [*This is not right*].

Three things follow. First that the nature hypothesis of the DH is fundamentally flawed. Second that the knowledge hypothesis of the DH is flawed. Third, and lastly, that the threat from nontrivial causal spread is equally unsound.

5.2. Nontrivial causal spread & representation-hungry problem-domain

The anti-representational (empirical) demonstrations used to motivate anti-representationalism fall within the class of purely reactive, situationally-determined behavior, thus pertaining only to systems that are uninterestingly non-cognitive. This is the second premise in the argument against anti-representationalism as a well-founded theory of cognition. In order not to beg questions, an additional argument will be introduced in support of this premise. Its logical form is:

- i. Behavioral autonomy and intentionality are necessary and sufficient conditions in order to demarcate genuinely cognitive systems from merely reactive systems.
- ii. A necessary prerequisite for behavioral autonomy and intentionality is the presence of internal states that “*stand-in*” for immediate perception, and function as “*control structures*” for action.

- iii. The anti-representational (empirical) demonstrations of intelligent, adaptive behavior are *unable* to meet these requirements. *Therefore:*
- iv. The anti-representational (empirical) demonstrations are merely reactive in character. *Therefore:*
- v. The anti-representational (empirical) demonstrations do not justify as genuinely cognitive systems.

5.2.1. *Stipulative definitions - intentionality & behavioral autonomy*

Intentionality, as the term is put to use here, specifies the *information-bearing* properties of certain neurophysiological processes or states in the brain of human and nonhuman animals that consists in their having the function of carrying specific types of *semantically evaluable content*. There is nothing controversial about this stipulation. But note that in the context of *dynamical cognitive science*, this account conceptualizes the “*aboutness*” of intentionality in terms of *possibilities for (inter)action* and that such action-oriented states have *correctness conditions* (see e.g., Hutto ,1999, p. 58; Toribio, 2007, p. 446). What this means is that intentional, action-specifying states are susceptible to error detection - i.e., such states can be detected as correct (true) or incorrect (false) *for the animal itself*. Note that this stipulation of intentionality corresponds to the definition of representation on the interactional and teleological accounts.

Behavioral autonomy, as the term is put to use here, specifies the ability of an agent to bring its *own experiences* onto a particular situation, and from which the agent’s action will emerge from the *joint interactions* of its experiences and the current contextual situation (see e.g., Beer, 1995, p. 173; Ziemke, 1999, p. 183). Note that an autonomous agent, in this *behavioral* sense, is fundamentally different from a *reactive agent*. According to Pfeifer, a reactive agent does not incorporate its own background experience – it always reacts the same way in the same situation (1995, p. 47-70). A second significant feature of behavioral autonomy has to do with *automaticity*; in fact, automaticity is a necessary element of any kind of autonomy (Ziemke, 1998, p. 568). By stipulating behavioral autonomy as *automatic*,

this means that for an agent (\mathcal{A}) to be behaviourally autonomous, (\mathcal{A}) must exhibit the ability to *adapt* its behavior within a dynamic environment, and (\mathcal{A}) must do so in ways that are *appropriate* (or beneficial) in current situations (see also Bickhard, 2004, p. 78). Note that this stipulation congers with the interactional and teleological accounts of representation.

5.3. *The flaw of the threat from nontrivial causal spread - infant stepping*

The basic trouble identified concerning the case of anti-representationalism based on nontrivial causal spread is that the sceptic substantially overstates his or her case. It seems to me that much of the scepticism stems from an unwarranted conflation of *extreme nontrivial causal spread* with cases of nontrivial causal spread *in general*. What lends support for this claim is that the demonstration of infant treadmill stepping is exactly such a conflation. It illustrates that the stepping behavior is *not* governed by neural mechanisms, but causally *spread* to factors *spatially distanced* from prior neural intuitions - in the non-neural body and environment. It is important to notice that this case study is set up so as to show that the presence of nontrivial causal spread can be used to drive an anti-representationalist assault if coupled with an additional claim: the claim that no *neural factors* can meet the *responsibility requirement* of internal representational accounts - in order for (\mathbf{R}) to justify as a representation of its target (\mathbf{T}), (\mathbf{R}) must be located in the central nervous system, (\mathbf{T}) must be fully absent, and (\mathbf{R}) must be *responsible* for adaptive, intelligent action. This is a case of *extreme nontrivial nontrivial causal spread*. Not only are additional cognitive processes *localized* beyond the bounds of the brain - in this case, the treadmill. In fact, these spatially extended processes are argued to be *the* constitutive processes. In so doing, it follows that the demonstration of infant treadmill stepping falls within the ontologically *extreme* class of purely reactive, situationally-determined activity on the continuum of organismic activity. By conflating cases of *extreme nontrivial causal spread* with nontrivial causal spread *in general*, the anti-representationalist ignores several substantial features. First, the conflation overlooks that in cases of nontrivial causal spread, where cognitive processes are both embodied and

situated, *interactivist representations* - teleological and interactional - may play a key role in promoting organismic adaptivity. Second, dismissing the possibility that real-time, intelligent action requires the presence of internal representations - even in a situated and embodied sense - *inhibits* explaining (intelligent) behavior in terms of intentionality and behavioral autonomy, and, therefore, *delimits* the anti-representational demonstrations from falling within the ontological class of genuinely cognitive systems. Support favouring an interactivist account of inner representation *coupled* with nontrivial causal spread is the case of *motor emulation*.

5.4. *Interactivist representation & motor emulation - a case for the second premise*

Skilled reaching is the smooth, trouble-free orientation of an arm and hand system directed at some target object. The achievement of real-time, adaptive success in cases of rapid on-line reaching depends, so Clark & Grush inform us, upon the CNS receiving and responding to proprioceptive feedback – in particular, when such information is not visually available (1999, p. 6). Proprioceptive feedback here involves information about arm/hand orientation, position and trajectory. What is interesting about skilled reaching behavior is that due to an inherent speed limitation of the CNS, proprioceptive feedback from extra-neural states (in this case, bodily) is often required *faster* than it is available. It is in such instances that *motor emulation* is crucial so as to enhance fluent, real-time action. Motor emulation is a piece of motor control circuitry, whose proper function is to *predict* sensory feedback prior to the signals arriving from the bodily peripheries (Clark, 1997b, p. 471). Hence, an emulator is a control system, within an overall system, that takes as *input* information concerning the current state of the system (e.g., the state of the arm, direction of motion, etc.) and gives as *output* a prediction of the future feedback from the arm/hand system, thereby dynamically presupposing future positioning and trajectory of the arm before actual feedback arrives. On my account, as well as Clark & Grush's, motor emulation is important, because it illuminates a set of features of significance with respect to internal representation.

First, the *output* of the emulator circuitry *stands-in* for immediate perception,

and functions as *control structures* for action. Earlier we identified these conditions as necessary in order to justify as a representation. In fact, the output of motor emulation is compatible with what I earlier termed the *minimal decoupleability* condition for *interactional representations*: (\mathbf{R}) – the output of emulation – and its target (\mathbf{T}) are *not* in constant causal contact, (\mathbf{T}) is *temporally* absent – the emulator is not receiving input directly from what it represents – and (\mathbf{R}) is adaptive – (\mathbf{R}) *enhances* behavioral success in the temporal absence of immediate stimuli. *Prima facie* motor emulation allows for an evolutionary minimal sense of *behavioral autonomy*: it is functionally disconnected from behavioral reactivity and situation determination, and it brings its own *embodied* experiences (i.e., dynamic predictions) to bear in determining action.

Second, the *content* carried by the inner vehicles of the emulation circuitry is *semantically evaluable* – it is subject to conditions of correctness. For instance, if the predictions do not hold – turn out to be correct – then the *content* of the predictions are *incorrect* (see also Bickhard, 2004, p. 78). Because of this, in the case of real-time, representation-hungry activity, the content of inner interactivist vehicles of content can be true (correct) or false (incorrect), and can be so about the environment. Motor emulation, so it would seem, constitutes the minimal requirements for intentionality and behavioral autonomy, which, in turn, are both necessary and sufficient conditions in order to demarcate genuinely cognitive systems from merely reactive systems.

5.5. *The first flaw of the nature & knowledge hypotheses of the DH*

Justifiability for an anti-representational account of cognition turns on the *completeness condition*: In order for a theory, (\mathbf{T}), to count as a well-founded anti-representationalist theory of cognition, (\mathbf{T}) must involve *no* use of representation-laden concepts, and (\mathbf{T}) must have conceptually *sufficient* means to explain cognitive phenomena *in toto*. This is the first premise of the argument against anti-representationalism as a well-founded theory of cognition. Why think that the completeness condition is questionable?

Some authors – especially, Chemero (2000) and Bechtel (1998) – question

van Gelder's (1995) claim that the WG is a nonrepresentational dynamical system. That is, these authors question the first horn of the first premise: that (**T**) must involve *no* use of representation-laden concepts. On Chemero's view, the WG can be interpreted such that the *angle of the arm* (representation) carries information about the *speed of the flywheel* (representation producer) and the *throttle valve* (representation consumer) uses the information about arm angles to coordinate its behavior, and so adapt to the represented speed of the flywheel (Chemero, 2000, p. 632). Because of this, one plausible route to take against the anti-representational sceptic is to make the *conceptual* claim that the arm angles of the WG count as internal representations *for* the throttle valve mechanism. Both of these authors therefore go on the claim that we have no reason to think that cognition is inherently nonrepresentational (see e.g., Bechtel, 1998, p. 299). Although this argument *prima facially* puts pressure on both the nature and the knowledge hypothesis of the DH, I do not find it to be a very attractive response to make against the sceptic. Two reasons support my position.

The first reason turns on the suspicion that *if* representations become ubiquitous, what extra value can they have for the study of cognition? This worry is exactly why I agree with the DH that there is a large class of dynamical models for which explanation in terms of representation has *no* utility. If the notion of representation is applicable to every system from purely reactive, situationally-determined systems such as the WG..., to highly cerebral, situation-independent systems, then there is a real danger that the notion of representation threatens to lose its explanatory substance. In fact, the value of *not* interpreting the WG as a representational system is that it allows us to demarcate representational systems from systems to which the notion does *not* apply. The second reason critically targets the second horn of premise one: That (**T**) must have conceptually *sufficient* means to explain cognitive phenomena *in toto*. Evidence for this critique is provided by premise two above: A nonrepresentational theory of cognition is unable to meet the minimal requirements for genuine cognition. Here the first flaw of the nature hypothesis and the knowledge hypothesis of the DH reveals itself:

On the one hand, ontologically, if cognitive agents *are* dynamical systems, and a minimum requirement for genuine cognition is the presence of interactionist vehicles standing-in for immediate perception, then *some* dynamical systems *are* representational using systems. It is just that the WG is not such a system. Therefore, one is not justified in inferring from the nonrepresentational nature of the WG to cognition in general (see e.g., Clark & Toribio, 1994, p. 421). This has implications for the nature hypothesis of the DH, because it suggests that *some* activities, performed by cognitive agents, *do* involve the presence of internal representations. On the other hand, epistemologically, if cognitive science ought to adopt a dynamical approach to the study of cognition, and *some* cognitive activities fall within the class of real-time, representation-hungry activity, then *some* dynamical explanations, even if these do not explicitly mention internal representations, must presuppose the use of internal representations by the systems under investigation. This has implications for the knowledge hypothesis of the DH. First, it suggests that the conceptual tools of DST are important with respect to dealing with complex, coupled dynamical systems – like the WG. Second, only if the operations of a dynamical system is driven by sufficient amounts of ambient stimuli, *then* the dynamical systems theoretical explanation needed will be nonrepresentational; otherwise not. Therefore, DST ought not sidestep representational issues altogether, but rather attempt to articulate representations within the dynamical framework itself (for related views see Bickhard & Terveen, 1995; Bickhard, 2003, 2004; Christensen & Hooker, 2004; Clark & Toribio, 1994; Clark, 1997a, 2001; Eliasmith, 1996, 1997; and others).

5.6. The second flaw of the nature & knowledge hypothesis of the DH

Cognitive properties are emergent properties, but emergent properties – together with the tools of dynamical systems theory – are not intrinsically nonrepresentational. This is the third premise offered in the dispute against anti-representationalism as a well-founded theory of cognition. Here we will confront what I, among others, take to be the hardest of the sceptical positions: Cognitive properties are emergent properties, and an emergent property, (*Sp*), depends on

continuous reciprocal causation among its subsystemic parts (*Sp1, Sp2, ..., Spn*), thus making it *empirically impossible* to assign specific functional and/or representational roles to *scientifically identifiable* subsystems (see e.g., van Gelder, 1995, p. 354; Wheeler, 2005b, p. 225). Generally speaking, the sceptics conclude, *if* a systemic property is emergent, *then* it defies modular explanation; modularity is necessary for internal representation; so, if continuous reciprocal causation prevents modular explanation, and modularity is necessary for representation, then to prevent modular explanation is also to resist representational explanation. My hypothesis is, *insofar* as the WG is meant to highlight the justifiability of a well-founded, anti-representational theory of cognition, it fails to do so. Not only because of the flaws identified thus far, but since it seems to me that a systemic property – for instance, change in arm angle in the WG – *is* an emergent property and (importantly) *without* being resistant to modular explanations.

Emergent properties are *systemic properties*. A property of a system is systemic *if and only if* it is a collective property of the overall system and not a property of any part of the system. The WG functions so as to maintain the speed of the flywheel constant in spite of fluctuations in steam pressure and workload. The systemic property of the WG is *constant speed*. Of special significance is that “constant speed” emerges from what Clark calls *emergence as unprogrammed functionality* (2001, p. 114): If (*Sp*) is an emergent property of a system, then (*Sp*) arises as a kind of unprogrammed “side-effect” from the interactions among the variables comprising the system (*Sp1, Sp2, ..., Spn*). For instance, the WG and the engine system constitutes a coupled dynamical system, in which θ (the angle of the governor’s arms) and w (engine speed) function as subsystemic variables. Because of this, constant speed is a collective and unprogrammed property, whose value reflects the interactions of θ and w comprising the system.

An emergent property is dependent on continuous reciprocal causation among its individual parts. Clark classifies this feature as *emergence from interactive complexity*. By this he means that “emergence [is] the process by which complex, cyclic interactions give rise to stable and salient patterns of systemic behavior.” (2001, p. 114; square bracketed quotation added). On the one hand, this entails

that there can be no difference in a systems systemic properties unless there is a difference in the properties of the system's parts or their mode of organization. Constant speed is thus directly dependent on θ and w . On the other hand, it illustrates that an emergent property is the result of circular, interdependent causal processes among its individual parts. So, constant speed is a function of w , which itself is a function of θ , and where any change in w changes the dynamics of θ – hence, constant speed is indirectly a function of θ .

The functioning of the WG is (interestingly) emergent – that is, we gain significantly by treating it as such and by providing a dynamical systems style of description to capture its overall dynamics. But for all that, *if* we know the variables of the system and once a specification of these variables has been given by a set of differential equations, *then* the behavior of the WG has been given a modular explanation as well. For example, the equation, ($dnw/dtn = f(w, \dots, t, \dots)$), specifies a set of individual variables – w *and* t – and explains their mode of organization – w is a function of current engine speed, which itself is dependent on t (the throttle valve setting), which is a function of θ – and it defines “coordinated coupling” as an emergent property. Therefore, the WG allows for an account of emergent behavior, together with an account that *identifies* the individual variables and their overall organization. Importantly, since van Gelder offers the WG as a paradigmatic dynamical system, and as a prototypical model for a new paradigm in cognitive science, the fact that the WG is not resistant to modularity opens a *space of possibilities*: It implies that other dynamical models of cognition – such as those that fall within the class of representation-hungry activities – might be emergent *plus* modular – modularity being a necessary condition for empirically identifying representations within a system. Therefore, neither dynamic systems nor DST are inherently anti-representational, and, consequently, anti-representationalism cannot be justified as a well-founded theory of cognition.

6. Conclusion

In the foregoing sections, I have presented arguments to show that the links

between the dynamic *demonstrations* – phonotaxis, infant stepping behavior, Watt governing, avoiding obstacles... – and the sceptical *conclusion* are too weak (i.e., insufficient) for the demonstrations to justify the radical anti-representational conclusion. Hence, I think it is justified - relative to the premises I have used and tried to substantiate - to conclude that anti-representationalism *per se* is an ill motivated theory of cognition. Parsimony, I have argued, rules in favour of rejecting the need for a representational analysis. This is a major strength of the anti-representational position – it makes evident the existence of the class of dynamical systems for which representational analysis has no scientific value. Nonetheless, we have found reasons upon which to base the claim that we need not accept the much stronger sceptical conclusion – namely, that the best explanations of cognitive phenomena fall within this class. Rather than denying the need for representation, I have tried to highlight that representational explanation – reconstructed within a dynamical agent-environment characterization – serves an indispensable explanatory task: It enables us to demarcate activities that presuppose intentionality and behavioral autonomy from activities that are merely reactive and situation- determined. The former falls within the class of representation-hungry activities, whereas the latter does not. Hence, anti-representationalism *per se* is an excessively costly undertaking, since it prevents cognitive science from dealing with vitally important issues such as intentionality and behavioral autonomy.

References

- Ballard, D. (1991). 'Animate Vision'. *Artificial Intelligence* 48: 57-86.
- Bechtel, W. (1998). 'Representations and Cognitive Explanations: Assessing the Dynamicist's Challenge in Cognitive Science'. *Cognitive Science* 22: 295-318.
- Beer, R. (1995). 'A Dynamical Systems Perspective on Agent-Environment Interaction'. *Artificial Intelligence* 72: 173-215.
- Bickhard, M., & Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and*

- Cognitive Science: Impasse and Solution*. Amsterdam: Elsevier Science Publishers.
- Bickhard, M. (2003). 'Process and Emergence: Normative Function and Representation'. *Axiomathes* 14: 121-155.
- Bickhard, M. (2004). 'The Dynamic Emergence of Representation'. In: Clapin, H., Staines, P., & Slezak, P. (Eds.) *Representation in Mind: New Approaches to Mental Representation*. Amsterdam: Elsevier B.V.
- Brooks, R.A. (1990). 'Elephants Don't Play Chess'. *Robotics and Autonomous Systems* 6: 3-15.
- Brooks, R.A. (1991). 'Intelligence Without Representation'. *Artificial Intelligence* 47: 139-59.
- Brooks, R.A. (1999). 'Intelligence Without Reason'. *Cambrian Intelligence*. Cambridge, Mass...: The MIT Press: 133-186.
- Chemero, A. (2000). 'Anti-Representationalism and the Dynamical Stance'. *Philosophy of Science* 67: 625-647.
- Chiel, H.J., & Beer, R.D. (1997). 'The Brain has a Body: Adaptive Behavior Emerges From Interaction of Nervous System, Body and Environment'. *Trends Neurosci* 20: 553-557.
- Christensen, W., & Hooker, A.C. (2004). 'Representation and the Meaning of Life'. In: Clapin, H., Staines, P., & Slezak, P. (Eds.) *Representation in Mind: New Approaches to Mental Representation*. Amsterdam: Elsevier B.V.
- Clark, A., & Toribio, J. (1994). 'Doing Without Representing?'. *Synthese* 101: 401-431.
- Clark, A. (1996). 'Happy Couplings: Emergence and Explanatory Interlock'. In: Boden, M. (ed.) *The Philosophy of Artificial Life*. New York and Oxford: Oxford University Press.
- Clark, A. (1997a). *Being There*. Cambridge, Mass...: The MIT Press.
- Clark, A. (1997b). 'The Dynamical Challenge'. *Cognitive Science* 21 (4): 461-481.
- Clark, A., & Grush, R. (1999). 'Towards a Cognitive Robotics'. *Adaptive Behavior* 7 (1): 5-16.
- Clark, A. (2001). *Mindware*. New York and Oxford: Oxford University Press.
- Costa, F., & Rocha, L.M. (2005). 'Introduction to the Special Issue: Embodied and Situated Cognition'. *Artificial Life* 11: 5-11.
- Damasio, A. (1994). *Descartes' Error*. London: Vintage Books.
- Damasio, A. (2000). *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. London: Vintage Books.
- Dorffner, G. (1999). 'The Connectionist Route to Embodiment and

- Dynamicism'. In: Riegler, A., Peschl, M., & von Stein, A. (Eds.) *Understanding Representation in the Cognitive Sciences: Does Representation Need Reality?* New York: Kluwer Academic/Plenum Publishers.
- Eliasmith, C. (1996). 'The Third Contender: A Critical Examination of the Dynamicist Theory of Cognition'. *Philosophical Psychology* 9: 441-463.
- Eliasmith, C. (1997). 'Computation and Dynamical Models of Mind'. *Minds and Machines* 7: 531-541.
- Fodor, J. (1993). *The Modularity of Mind*. Cambridge, Mass...: The MIT Press.
- Fodor, J., & Pylyshyn, Z. (1988). 'Connectionism and Cognitive Architecture: A Critical Analysis'. *Cognition* 28: 3-71.
- Freeman, W.J. (1999). *How Brain Make Up Their Minds*. London: Phoenix.
- Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. New Jersey: Lawrence Erlbaum Associates.
- Haselager, P., Bongers, R., & van Rooij, I. (2003a). 'Cognitive Science, Representations and Dynamical Systems Theory'. In: Tschacher, W., & Dauwalder, P.J. (Eds.) *The Dynamical Systems Approach to Cognition: Concepts and Empirical Paradigms Based on Self-Organization, Embodiment and Coordination Dynamics*. New Jersey & London: World Scientific Publishing.
- Haselager, P., de Groot, A., & van Rappard, H. (2003b). 'Representationalism vs. Anti-representationalism: A Debate for the Sake of Appearance'. *Philosophical Psychology* 16 (1): 5-24.
- Haugeland, J. (1991). 'Semantic Engines: An Introduction to Mind Design'. In: Haugeland, J. (ed.) *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge, Mass...: The MIT Press.
- Heidegger, M. (1927/2001). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag GmbH.
- Hendriks-Jansen, H. (1996). 'In Praise of Interactive Emergence, or Why Explanations don't have to wait for Implementation.' In: Boden, M. (ed.) *The Philosophy of Artificial Life*. New York and Oxford: Oxford University Press.
- Keijzer, A.F. (1998). 'Doing Without Representations Which Specify What To Do'. *Philosophical Psychology* 11 (3): 269-302.
- Kelso, S. (2003). 'Cognitive Coordination Dynamics'. In: Tschacher, W., & Dauwalder, P.J. (Eds.) *The Dynamical Systems Approach to Cognition: Concepts and Empirical Paradigms Based on Self-Organization, Embodiment and Coordination Dynamics*. New Jersey & London: World Scientific Publishing.
- Hutto, D.D. (1999). 'Cognition Without Representation'. In: Riegler, A., Peschl, M., & von Stein, A. (Eds.) *Understanding Representation in the Cognitive Sciences:*

- Does Representation Need Reality?* New York: Kluwer Academic/Plenum Publishers.
- Merleau-Ponty, M. (1945/2002). *Phenomenology of Perception*. London and New York: Routledge.
- Kirsh, D. (1996). 'Today The Earwig, Tomorrow Man'. In: Boden, M. (ed.) *The Philosophy of Artificial Life*. New York and Oxford: Oxford University Press.
- Millikan, R. G. (2004). *Varieties of Meaning*. Cambridge, Mass...: The MIT Press.
- Newell, A. (1980). 'Physical Symbol Systems'. *Cognitive Science* 4: 135-183.
- Peschl, M.F., & Riegler, A. (1999). 'Does Representation Need Reality? Rethinking Epistemological Issues in the Light of Recent Developments and Concepts in Cognitive Science'. In: Riegler, A., Peschl, M., & von Stein, A. (Eds.) *Understanding Representation in the Cognitive Sciences: Does Representation Need Reality?* New York: Kluwer Academic/Plenum Publishers.
- Pinker, S. (1997). *How the Mind Works*. London: Penguin Books.
- Smith, B.C. (1996). *On the Origin of Objects*. Cambridge, Mass...: The MIT Press.
- Stephan, A. (2006). 'The Dual of 'Emergence' in the Philosophy of Mind and in Cognitive Science'. *Synthese* 151: 485-498.
- Sterelny, K. (1990). *The Representational Theory of Mind*. Cambridge, Mass...: Basil Blackwell.
- Thelen, E., and Smith, B.L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Mass...: The MIT Press.
- Thelen, E. (2003). 'Grounded in the World: Developmental Origins of the Embodied Mind'. In: Tschacher, W., & Dauwalder, P.J. (Eds) *The Dynamical Systems Approach to Cognition: Concepts and Empirical Paradigms Based on Self-Organization, Embodiment and Coordination Dynamics*. New Jersey & London: World Scientific Publishing.
- Toribio, J. (2007). 'Nonconceptual Content'. *Philosophical Compass* 2 (3): 445-460.
- Van Gelder, T. (1995). 'What Might Cognition Be, If Not Computational?' *The Journal of Philosophy* 92 (7): 345-381.
- Van Gelder, T. (1998). 'The Dynamical Hypothesis in Cognitive Science'. *Behavioral and Brain Sciences* 21: 615-665.
- Webb, B. (1994). 'Robotics Experiments in Cricket Phonotaxis'. In: Cliff, D., Husbands, P., Meyer, J.A., & Wilson, S. (Eds.) *From Animals to Animats 3*. Cambridge, Mass...: The MIT Press.
- Webb, B. (1996). 'A Cricket Robot'. *Scientific American* 275: 62-67.
- Wheeler, M. (1994). 'From Activation to Activity: Representation, Computation, and the Dynamics of Neural Network Control Systems'. *AISB Quarterly* 87:

36-42.

- Wheeler, M. (1996). 'From Robots to Rothko: the Bringing Forth of Worlds'. In: Boden, M. (ed.) *The Philosophy of Artificial Life*. New York and Oxford: Oxford University Press.
- Wheeler, M., & Clark, A. (1999). 'Genic Representation: Reconciling Content and Causal Complexity'. *British Journal for the Philosophy of Science* 51 (1): 103-135.
- Wheeler, M. (2001). 'Two Threats to Representation'. *Synthese* 129: 211-231.
- Wheeler, M. (2005a). 'Friends Reunited: Evolutionary Robotics and Representational Explanations'. *Artificial Life* 11: 215-31.
- Wheeler, M. (2005b). *Reconstructing the Cognitive World*. Cambridge, Mass.: The MIT Press.
- Ziemke, T. (1998). 'Adaptive Behavior in Autonomous Agents'. *Presence* 7 (6): 564-587.
- Ziemke, T. (1999). 'Rethinking Grounding'. In: Riegler, A., Peschl, M., & von Stein, A. (Eds.) *Understanding Representation in the Cognitive Sciences: Does Representation Need Reality?* New York: Kluwer Academic/Plenum Publishers.