Towards a general theory of antirepresentationalism*

To appear in The British Journal for the Philosophy of Science

Francisco Calvo Garzón

fjcalvo@um.es

Abstract This work represents an attempt to stake out the landscape for dynamicism based on a radical dismissal of the information-processing paradigm that dominates the philosophy of cognitive science. In section 2, after setting up the basic toolkit of a theory of minimal representationalism, I introduce the central tenets of dynamic systems theory (DST) by discussing recent research in the dynamics of embodiment (Thelen et al., 2001) in the perseverative-reaching literature. A recent proposal on the dynamics of representation—the dynamic field approach (Spencer & Schöner, 2003)—according to which the alleged representational gap between DST and representational theories of cognition needs to be bridged in order to explain higher-order cognitive activity will then be reviewed. In section 3 I shall argue that Spencer and Schöner’s attempt to bridge the representational gap may jeopardize the whole (antirepresentationalist) spirit of the DST project. In order to show why, I shall introduce the key concepts of “reliability of environment” and “primagenesis”, and argue that DST can account for de-coupled, offline cognitive activity with no need of positing representational resources. Conclusions and directions for future research will follow.

* I would like to thank Istvan Berkeley, Mark Bickhard, John Bickle, Richard Campbell, Jim Edwards, Toni Gomila, Stevan Harnad, Alicia Juarrero, Aarre Laakso, Gregor Schöner and John Symons for their comments on a previous version of this manuscript. I am also grateful to two anonymous referees for BJPS for helpful comments and suggestions. Preparation of this manuscript was supported by DGICYT Project BFF2003-129, and by a Ramón y Cajal Research Contract (Spanish Ministry of Science and Education).
Towards a general theory of antirepresentationalism

1. INTRODUCTION

The notions of computation and representation are not just common currency in cognitive science modelling. To put it mildly, they are the building blocks of the discipline. Alternative voices from a number of subdisciplines that call into question these notions have periodically been raised. Unfortunately, after an initial, and usually short, excitement they remain quiet. Silence is due mainly to two reasons: On the one hand, the dominant paradigm overwhelms competitors (sometimes due to “pragmatic” considerations) with data already accounted for and results to be accounted for, and on the other hand, alternative framings are repeatedly absorbed and made innocuous. Both reasons are interrelated. Alternatives raised, by default, carry the burden of proof in such a way that the dominant paradigm is the one that chooses what phenomena are in need of explanation. So far no one would complain. After all that’s the history of scientific progress, and as a matter of fact we’ve made some astonishing progress over the last few centuries! Problems start when the what limits the range of options available when it comes to answering the how. Ways of explaining cognition range from straightforward classical theories (Newell and Simon, 1972; Pylyshyn, 1984) to connectionist neural networks (O’Reilly and Munakata, 2000). Unfortunately, when frameworks that reject classical, connectionist and/or hybrid explanations are put forward (e.g., dynamic systems theory or situated cognition), the explananda posited by orthodoxy turns alternative explanans—e.g., non-classical computationalism (Scheutz, 2002); distributed representationalism (van Gelder, 1991)—into forms which, although interesting and empirically fruitful, lead models to collapse into mere variations of the same common theme. Namely, (did you guess it?) more computationalism and more representationalism. This work represents an attempt to stake out the landscape for dynamicism based on a radical dismissal of the information-processing paradigm that dominates the philosophy of cognitive science.¹

¹ Special caution is needed at these opening stages. Someone might argue that this introductory setting misrepresents the conceptual connections between dynamic systems theory and antirepresentationalism (thanks to an anonymous referee for BJPS for stressing this point). By favouring a DST perspective one is not thereby obliged to becoming antirepresentationalist. As a matter of fact, well-known foes of representationalism (e.g., van Gelder, 1995) have given up antirepresentationalism while remaining hard-core dynamicists (van Gelder, 1998). However, although focusing on antirepresentationalism from a dynamicist perspective certainly shows my personal biases, let me stress from the start that: (i) The sort of dynamicism that underlies my approach critically inherits the commitment to embodiment and embeddedness that underlied dynamicism in its pre-94 (Thelen and Smith, 1994) origins; and that (ii) a rebuttal of representationalism is not necessarily dependent upon agreement on the specific weight that DST by itself plays. For a comparison of a number of theories of development that distinguish between anti-representational (pre-94) dynamicism, and representational (post-94) dynamicism, see Thelen and Bates (2003).
1.1 Minimal representationalism

The information-processing paradigm posits sets of internal mechanisms that serve the purpose of information manipulation and storage. Although many dynamicists (e.g., Giunti, 1997) would insist on a somewhat more stringent notion of computation, for present purposes the key constraint is that the processes that govern input/output transformations are effective, that is, that we can find a set of instructions which is finite and allows us to obtain the output domain from the input. What’s at stake then in defining computationalism is the fact that effective computations can be performed over representations. Cognitive activity is thus marked by the processing of representational states. An information-processing agent counts as a computational system insofar as its state-transitions can be accounted for in terms of manipulations on representations. Put bluntly, the relation of representation refers to the standing in of internal states of a physical system for the content of other internal or external states. But, what is a representation, anyway? What does the “standing in” relation consist of? We can highlight some of the key features that any successful notion of representation presumably ought to incorporate. Dietrich and Markman (2003) provide a general definition of mental representation that intends to satisfy (almost) all parties to the debate. According to them,

a representation is any internal state that mediates or plays a mediating role between a system’s inputs and outputs in virtue of that state’s semantic content. We define semantic content in terms of information causally responsible for the state, and in terms of the use to which that information is put. (Ibid., p. 97)

And they continue,

Hence, any state that mediates between the input to an agent and the actions performed by the agent is a representation (this is somewhat unintuitive because it is so inclusive, but it has enormous benefits, especially in unifying cognitive science). (Ibid., p. 97)

The careful reader may have noticed that these two quotations are not equivalent. According to the second one, a representational state is any internal mediating state. However, the first quotation further qualifies this definition by means of a supplementary condition. The internal state in question must play the mediating role in virtue of the informational content to be associated with that state. In this way, Dietrich and Markman’s characterization amounts to the establishment of a precondition for a state to be representational. Namely, the satisfaction of what I shall dub the connectivity principle; a principle according to which for a state to be contentful it must play a critical role in the establishment of a connection between the system’s input and output states. For the role of such a connecting state to be critical, realism must be the default theoretical stance. When we claim that representational states have causal powers we mean to say that a system contains physical states that stand for other states and that can play a
causal role in the behavior of the system because of the content of the state in question. Crucially, representational states are causally efficacious states, in such a way that the system can behave adaptively by exploiting its informational states. That is, representations are for the system. For some authors (e.g., Bechtel, 1998), the state must have the function of “carrying the information embedded in the causal correlation.” Other authors (e.g., van Gelder, 1995) emphasize the explanatory contribution that needs to be made before a physical state can properly be called representational. For present purposes I shall simply grant a standard minimal (weak?) definition of representation, as the one implied by the connectivity principle under the lens of realism.

Someone may wish to argue that the connectivity principle effectively beefs up an important number of “go-between” states. Nevertheless, Dietrich and Markman point out the convenience of such an all-embracing conception since it renders unification in the cognitive sciences feasible. The question then, in view of these remarks, is: What else should be taken into consideration so that cognition doesn’t spread all over the board? In what follows, I shall introduce three further principles that taken together constitute a tripartite definition of representation.

Firstly, let me introduce what I shall dub the principle of minimal lastingness. According to this principle, for a physical state that satisfies the connectivity principle to become representational it must exhibit a degree of persistence that at least evens up with the degree shown by the system’s inputs to be represented. Notice that the emphasis is not on the fact that representational states are enduring states, but rather on the fact that typically they must endure longer (at this point I must ask the reader to grant the principle for argument’s sake. I shall motivate its adoption in section 3, below). Secondly, I shall consider a principle of dissociation according to which spatiotemporal unbinding is the mark of contentful states. For a state to count as representational for the system, the state must be able on occasions to stand for things or events that are simply lacking or temporarily unavailable. And finally, a principle of reification, which would follow naturally from a realist reading of the aforementioned connectivity principle. According to the principle of reification, a system state can only count

---

2 The connectivity principle by itself amounts to only a preliminary understanding of the concept of representation, insofar as it identifies a contentful state with a mere “go-between” state. Certainly, however, not all types of “go-between” states will be effectively beefed up into representations. Plausibly, there will be certain sorts of states which, though may well play a causal role between input and output, enjoy a metaphysically weightless informational interpretation. However, this may also be interpreted as a virtueless form of (at least, partial) panrepresentationalism, with the potential risk of saving representationalism nominally at the expense of stripping it of practical significance. For present purposes, I shall leave these matters aside (see Calvo Garzón for an elaboration of this issue).

3 Although related to well-known approaches such as Haugeland’s (1991) or Clark’s (1997), this principle (as well as the principles of dissociation and reification below) contains fine-grained differences that I shall not consider for present purposes (see Calvo Garzón).

as representational if it can be detected and a parallel drawn between the state in question and
the role it plays in the establishment of a connection between the system’s input and output
states. That is, we must be able to identify specific physical states with the (semantic) roles they
play (see section 3, below).

This tripartite definition, together with the connectivity principle, constitutes what I shall
call a theory of minimal representationalism (MR). In fairness to the truth, the set of principles
that gives rise to MR is to some extent unspecified. However, I only intend to characterize
what would be minimally required for a state to be representational, while keeping at a distance
from a pernicious degree of theoretical inclusiveness. MR claims that a physical state is
contentful if it can be spatiotemporarily identified as causally efficacious in the connection of
the system’s input and output states in such a way that the state in question “hangs in there”
while the input state it is tuned to decays or is no longer present. The reader will surely have in
mind other constraining principles that can be superadded. However, those other principles,
whatever their flesh is, are for the consideration of the sympathiser of representationalism. The
skeptic’s objective is to show that the physical states of a dynamical system fail to satisfy MR
in one form or another. That will suffice for my purposes. Whatever else is satisfied or not
becomes thus irrelevant.

Critics of orthodoxy argue from different fronts that the computational view of the mind
should be informed by non-classical research. Dietrich and Markman (2003) are well aware of
this point, but they claim:

The main way to attack computationalism, at least the main way it has been
attacked historically, is to show that cognition is incompatible with the notion of
representation. Logically speaking, there are two ways to do this. One can try to
show that cognition doesn’t require any representation at all, and that
representations just confuse the issues; or one can try to show that cognition does
require representations but that the representations have properties that prevent
them from being processed by algorithms. Despite the rhetoric (e.g., the title of
Brooks’ 1991 paper ‘Intelligence without Representation’) very few researchers
have tried the first approach. (2003, p. 96).

Hopes are that corrections in the scaffolding of the paradigm would make it more stable. My
view, however, is that a different scaffolding needs to be secured on firmer foundations. This
article is thus an attempt to go for the first thorny horn in Dietrich and Markman’s passage.

2. THE DYNAMIC FIELD APPROACH
Dynamic systems theory (hereafter, abbreviated DST) (Thelen and Smith, 1994) has been
gaining sympathisers in a number of disciplines in recent years. It is at the heart of disciplines

---

5 MR’s basic toolkit may be enlarged. So some authors emphasize the fact that different representations
may be modally linked differently, and that some may even be perceptually amodal. There are other
features of representations that I’ll ignore for present purposes (see Markman and Dietrich, 2000b).
such as synergetics, ecological psychology, and plays a key role in areas such as motor control, perception, development, and situated robotics. Bickhard & Terveen (1995), Clark (1997), Kelso (1995), and van Gelder (1998), to name but a few, are authors that in one form or another have embraced the theoretical approach of DST. The driving force of this joint effort is a distrust for cognitive science orthodoxy, in particular for its reliance on the foundational notions of computation and representation (although see fn. 1, above), and the belief that cognition ought to be studied in the continuous interactions of brain, body and environment.6

2.1 Dynamic systems theory and the continuity hypothesis

To fully appraise DST a distinction must be established between concrete dynamical systems and those abstract dynamical systems that the former may be said to instantiate. Concrete dynamical systems are real objects that change over time in the physical world. Abstract dynamical systems, on the other hand, are mathematical models that we can exploit to describe formally the actual changes undergone by concrete entities. An abstract dynamical system consists of a set of (non-linear) differential equations, where specific values of a subset of control variables amount to the state of the system in progress. Galilean mechanical objects are a classical example of concrete dynamical systems that are said to instantiate the laws that determine the position and velocity of falling bodies in space and time. The latter would specify a particular abstract dynamical system (Giunti, 1997). DST tries to model and explain the behavior of concrete systems by identifying them with sets of variables that change continually over time. A dynamical system, in this way, can be analyzed in terms of the differential equations that contain the quantitative variables whose interdependencies describe the laws that govern the behavior of the system.

One crucial aspect of DST is that changes in concrete dynamical systems are obtained continuously in time (see section 3 below). The other key concept is stability. DST defines stability as the endurance of a behavioral and/or neural condition under a certain form of perturbation which may range from complete systematicity to absolute randomness. The space delimited by the potential behavioral/neural conditions that the system may undergo is quantified by a set of (state) variables. In this way, given a possible condition (i.e., given certain values for the variables), the direction and rate of evolution of the states that the system can be in may be predicted. The particular values of the variables at which a vector predicts the absence of change is called a stable state. The system is said to converge to this sort of states from nearby points in phase space. There are a number of reasons why a system can achieve stability. It may be due to the external conditions of embodiment of the system, or for example, to endogenous neurally-generated and feedback-driven activity. Complex behavior, however,

---

6 Port and van Gelder (1995) is a good entry point to readers unfamiliar with the DST literature.
requires the generation of instabilities so that the system can evolve and reach new stable states. By *instability*, DST refers to the possible changes that can take a system from a stable state to an unstable one.\(^7\)

Now, if cognitive agents can be seen as concrete dynamical systems that instantiate mathematical models, DST may be interpreted as endorsing what I shall dub the “(cognitive) continuity hypothesis” (CCH). I shall distinguish two constraints that will serve to characterize the hypothesis. On the one hand, CCH, as an open empirical hypothesis on the architecture of cognition, incorporates a *homogeneity* constraint according to which (i) higher-level cognition is on a par with bodily and nervous processes, and can be thus accounted for in terms of continuity, temporality and non-linearity;\(^8\) and, on the other hand, it incorporates an *interactivity* constraint that says that (ii) multiple level interactions take place in a continuous fashion all the way down from high-level cognitive activity to perception, motor control and non-cognitive processes in general. In this way, stability and continuity in DST seem to go hand in hand. Both homogeneity and interactivity make of continuity one of the pivotal concepts in explaining behavior and cognition. In order to evolve continuously towards stability, variables in a dynamic system make use of gradually changing information that the system can access from different sources at any point in time.

van Gelder (1998) distinguishes two major components of dynamicism: A *knowledge* hypothesis according to which cognition should be appraised dynamically. It simply commits you to the view that cognition may best be modelled by means of an abstract set of dynamical systems. And on the other hand, a *nature* hypothesis, according to which cognitive agents are concrete dynamical systems. CCH above calls primarily for the ontological commitment of the latter hypothesis, although by no means it excludes the former. Ultimately, we may say that a dynamical model is sound if it describes realistically certain dynamical aspects of the concrete system it tries to model. For present purposes, continuity will be the crucial property of concrete systems to keep track of.

### 2.2 The dynamic field approach: Bridging the representational gap?

Recently, a number of DST researchers have identified in the concept of continuity a hurdle for dynamicism when issues of representation arise in the context of higher-level cognitive activity. As Spencer and Schöner (S&S) (2003) point out, those system states which are subject to a representational treatment need not evolve in a continuous fashion. In short, representational

---

\(^7\) Strictly speaking, attractor states and the instabilities that obtain constitute the core functionality that makes dynamical systems *behavioral* in the first place. Examples, therefore, of Galilean friction-less dynamics don’t serve to fully illustrate the foundational property of dynamical systems. However, for present purposes, this non-technical introduction will suffice. I thank Gregor Schöner for bringing this point to my attention.

\(^8\) It is worth noting, for clarity of exposition, that ‘continuity’ has often been brought into play in the literature (e.g., Clark, 2001) to refer to something close to what I here mean by ‘homogeneity’.
states are not subject to CCH above. There is, put bluntly, a representational gap that marks a watershed between lower-level behavior that meets CCH and higher-level behavior that doesn’t; a gap that in view of S&S needs to be bridged. S&S offer as an illustration of the latter sort of case the following example:

Consider three situations where a person decides to grab a coffee cup on a table. In one case, the cup is clearly visible and far from other objects; in another case, the cup is surrounded by identical “distracting” cups; in a third case, vision of the cup is obstructed by a stack of journal articles. Assume, further, that the person makes an identical movement in these cases—an accurate, stable, efficient reach that successfully makes contact with the cup. Does the motor approach capture everything about this situation? The answer is “no”. This approach fails to capture differences in the representational states underlying these movements. In particular, in one case, there was a high degree of certainty regarding where to move, in the second case, the decision of where to move was much less certain, and in the third case, the decision took on a new flavor—rather than being generated based on visible information, the decision was generated based on a longer-term memory of the cup’s location. (S&S, 2003, p. 398; emphasis added).9

It is noteworthy that the concept of continuity does not establish a black/white dichotomy. As we move away from lower level on-line activity towards more complex de-coupled tasks, continuity starts to dissolve. Highly abstract off-line reasoning, for instance, may be said to retain a minimal amount of continuities. It seems then that any model that ultimately succeeds in accounting dynamically for, say, the coffee cups example, will have to incorporate the increasingly discontinuous character of representational states.

S&S introduce the Dynamic Field Approach (henceforth, abbreviated DFA). Its goal is to bridge the aforementioned gap by extending DST to incorporate representational states dynamics (Erlhagen & Schöner, 2002). The basic working hypothesis in S&S’s overall approach may be stated as follows: Whatever model ultimately succeeds in accounting in dynamic terms for higher-level cognitive activity will have to incorporate the inherently discontinuous character of representational states. S&S make use of the mathematical concept of activation as employed, for example, in connectionist theory (Calvo Garzón, 2003), where contentful states are equated with points in a geometrically characterized hyperspace. In like

---

9 An anonymous referee worries that the example S&S use for purposes of illustration amounts to a mere thought-experiment of the sort that cognitive scientists usually turn a blind eye to. For one thing, the assumption that a person makes an identical movement in all three scenarios would need to be substantiated. As the referee points out, it is certainly possible to detect differences in movement across all three situations (for example, by means of electromyograms or biophysical analyses of limb movement). This is a serious worry that may well jeopardize the very point S&S’s example aims to illustrate, and a thorough recast of the three-cup example may ultimately be required. In principle, however, macroscopic variables may provide a complete characterization of the system, in such a way that microscopic ones might be uniquely determined as a function of the macroscopic level. The worry may then be addressed mathematically (Schöner, personal communication) by reducing the problem space, for example, by means of the Center Manifold Theorem. Unfortunately, considering potential manageable extensions of S&S’s illustration along these lines, would take us far afield, and, for argument’s sake, I shall grant the thought-experiment scenario as originally presented by S&S.
vein, DFA fleshes out the concept of activation in terms of a topological or metric activation field. DFA conceptualizes physical systems by providing a geometrical space where possible activation states are studied in terms of their positions, trajectories, basins of attraction, etc. The emphasis is on the change from one state into another, rather than on activation states per se. In this way, activations in the coffee cup example may be plotted on a continuous dimension of reaching locations in space, with peaks of activation in the field—roughly equivalent to Churchland’s (1989) connectionist “hot spots”—indicating the existence of a coffee cup at a given position.

It must be noted, however, that the inflection point that takes us from a non-representational model to S&S’s representational dynamics is not the concept of an activation field in itself, which is interpretable in terms of ecologically derived “directional movement parameters” (Schöner et al., 1997; Erlhagen & Schöner 2002). The point of departure from “standard” non-representational DST stems from S&S’s emphasis on the concept of discontinuity:

The third key feature of representational states is that they can be discontinuous in content. There are times (particularly with adults) when the content of a response doesn’t appear to be systematically related to the content of previous responses. For instance, in a target detection task, the representational state that underlies the verbal response “the far left target” does not obligatorily evolve through intervening locations en route to the response “the far right target”. (S&S, 2003, p. 399 – Emphasis in original)

The target detection task in the above passage refers to the literature in infants’ perseverative reaching behaviour, a topic of research where DFA has already delivered notable benefits. Let us make DFA’s point of departure in terms of discontinuities clearer by briefly reviewing that line of research (see also section 3 below).

Thelen et al. (2001) go exhaustively over the literature on the well-known, although still highly controversial, “A-not-B error”, and offer a non-classical explanation of motor control and development in that context. In addition, Thelen et al. speculate as to how higher-level phenomena may be dynamically modelled. Although their emphasis is on the inherently embodied character of cognition, an appraisal of their main tenets will be crucial to understand more abstract (and in principle, less-embodied) cognitive phenomena, which is the target of this paper.

The careful reader may have noticed the shift from “decoupability” to “discontinuity” between the last quote and this coming one. In addition to discontinuity, S&S highlight two more key features: graded certainty and bi-stability that provide the mathematical toolkit to establish the connection. These features relate to the alleged sufficiency of available information coded in activation space (e.g., about the location of the coffee cup) required to generate the appropriate movement (see S&S, 2003, for the details). However, for present purposes, we may ignore the fine-grained detail and focus exclusively on the concept of discontinuity, granting that where one has decoupled activity, one is not thereby necessarily committed to continuous activity (see section 3.3, below).
The original version of the A-not-B error (first discussed by Piaget, 1954) can be stated very succinctly. 7-to-12 month old infants who manage to unveil an object situated at an A location keep reaching to that area despite having clearly seen that the object has been removed and hidden at a different B location. The orthodox interpretation of this phenomenon in the developmental literature tells that the error observed has to do with the stage in concept acquisition the infant is in. In Piaget’s terminology, the infant is said to be at this point at “stage IV”. At this stage, the infant only credits the object with lasting existence at the location it has been seen appearing. The error is a cognitive one, and is explained in terms of the gradual development of the infant’s conceptual apparatus. Thelen et al. resist the conclusion that the error made shows the acquisition of concepts in infants. Crucially, Thelen et al. note, small variations in the experimental conditions can upset the robustness of the canonical error form. All visual properties, different time delays between hiding and searching, existence of landmarks, type of object being searched for, previous crawling experience, etc., etc., can make a difference to the final outcome. The moral is that no single factor is responsible for the effects observed. The bad news for the orthodox interpretation of the error is that the representation of the object that the infant searches for, if it is to be classically appraised, cannot possibly depend on such a wide variety of contextual factors.

In their view, the A-not-B error can be perfectly explained in terms of a four-fold coupled dynamics integrated by the actions of looking, planning, reaching, and remembering, with no need of invoking classical mentalistic concepts and operations. Specifically, in order to understand the multiple interactions that constitute these dynamic couplings, Thelen et al. focused on the reaching process that leads infants either to A or B. While reaching for the object, the target position must be retained. In addition, long-term memory of previous reaches biases subsequent ones. Crucially for the dynamic explanation, this amounts to a “preshaping” of the activation field. Their working hypothesis is that the error is due to motor planning in the context of perception and remembering.

Figure 1 shows a simulation of an infant’s perseverative reaching behavior. Dynamic fields (right columns) show range of reaching locations along the abscissa. Left columns show particular inputs being fed in the simulated task. Peaks of activation (fig. 1a) represent hiding locations A and B in the canonical form of the error. Fig 1b represents the process of hiding the toy (specific input). Figure 1 shows how inputs mesh in the working memory field. Before reaching there’s higher activation at the A location. As the object is waved at B, activation at the B location augments. Subsequently, as the object is hidden, there is a decay of activation at B. Due to the memory effect, when the infant finally reaches for the object, the higher level of activation at A exercises control over the behavior, triggering the pattern of error.

Notice, by contrast (fig. 1, bottom) that 12-month-old infants don’t make the error anymore. Although the input values coincide with the top panel ones, evolution in working
memory field differs. As the object is hidden, there is no decay of activation at B, remembering thus the act of hiding the object at the B location. This is due to the different way in which activation excites/inhibits neighboring/further away sites. In the case of older infants, endogeneous interactivity reverberates by means of inhibition/excitation connections, making particular inputs to the system have less weight with regard to the overall output behavior (see S&S, 2003, for the details, and section 3, below).

The dynamical evolution of the coupling of perception, movement and memory explains by itself the A-not-B error. DFA accounts for deviations from the canonical form of the experiment by means of a motor field where evolution of the system can be studied as all factors come into play. Crucially, cognitive activity cannot be accounted for without taking into the perceptual and motor apparatus that facilitates in the first place the agent’s dealing with the external world. DST sees embodiment (Thelen and Smith, 1994; Clark, 1997) as the mark of the mental. In this way, learning, memory, attention, etc., stand in sharp contrast to the classical viewpoint, according to which mental activity is ultimately cashed out in terms of symbol-manipulation. By contrast, a model of mental activity must respect the same principles of non-linearity, time-dependence and continuity that are generally invoked in explanations of bodily interactions and neural activity (Freeman, 2000).

Discontinuities in content find a natural interpretation in DFA. A change of representational content is described in terms of the evolution of states of activation in the field. As S&S remark, in the case where an environmental target is lost to sight and another one enters the picture a specific pattern of activation will indicate that change. Activation linked to the first object will decay while activation related to the second object will increase. As S&S point out, “Importantly, this can occur even though activation never builds up at intermediate reachable locations” (p. 399). So, DFA incorporates what I shall dub the thesis of content discontinuity (CD). CD says that for a cognitive system to count as re-presentational it is sufficient, although not necessary, that activation states can be discontinuous. Two states are discontinuous when field evolution does not require the activation of all locations in between the states to build up. Notice that endorsing CD apparently does not go against the aforementioned (cognitive) continuity hypothesis (CCH):

\[\text{INSERT FIGURE 1 ABOUT HERE}\]

\[\text{11 It may be argued that the long-term memory of previous reaches constitutes a representation-based resource, where memory-related activation states guide the movement through some sort of explicit information recall. As we shall see in due course such a rendering would result in a fatal misinterpretation of the dynamicist approach.}\]
This is not to say that the processes underlying these two representational states were not continuous, nor does it imply that continuous change of representational state never occurs (e.g., continuous state change does occur in many mental rotation tasks). The central point is that the content of representational states does not necessarily take on this continuous character. This stands in contrast to the motor system which must always evolve continuously in “content”. Thus, a pointing response to the “left” must necessarily evolve through intervening locations en route to the “right”. (ibid., p. 399)

The coming arguments will serve to further specify the critical relation between CD and CCH.

3. TOWARDS A GENERAL THEORY OF ANTIREPRESENTATIONALISM

In this section I shall argue that adaptive behaviour in an ecological setting, such as the type of behaviour humans exhibit in the world, need not rely on the exploitation of internal representations. As a reminder, internal representations are understood as physical states of a system that meet the demands of the tripartite definition I offered in section 1.1 above, which together with the connectivity principle, constitute what I called a theory of minimal representationalism (MR). In this way, the objective of this section is to show that the physical states of a dynamical system fail to satisfy MR. As we move towards higher level phenomena, and away from sensory and motor origins, the tendency is to endorse a representationalist stance. However, the notions of reliability of environment and primagenesis (3.2, below) will help us appraise why the dynamical states in question do not “hang in there” while the input states they are tuned to decay or are no longer present, failing thus to meet MR’s principle of minimal lastingness. In my view, it is not clear that DFA ought to be interpreted representationally, or at least we should motivate doing so.

3.1 Diagonal systems and microstimulated dissociations

As the notion of computation shrinks, friends of dynamicism and representationalism may feel tempted to still have it both ways, and opt for a “soft” notion of representation and/or computation (e.g., Wheeler and Clark, 1999; Clark & Grush, 1999). According to a dynamical “soft” form of representation, no inferences that involve abstract items beyond the contextualized patterns of behaviour of the system should be posited. But, can we question the very nature of representations, independently of whether they are “soft” or “hard”? That is, can we question the thesis of minimal representationalism? If we are to trespass that (minimal) representational line, a great deal will depend on realizing that the following picture does not exhaust the realm of possibilities:

Certainly we accept the bare possibility that biological brains might constitute systems capable of de-coupled, off-line modeling yet not prove susceptible to the

---

12 Clark and Toribio’s (1994) appeal for “representation-hungry cases” is an illustration of this tendency.
kind of functional de-composition involved in a representational analysis. This might be the case if, for example, the inner goings-on are all interlinked by vast quantities of reciprocal causal influence, thus making it impossible to assign specific representational roles to identifiable sub-states or processes. Under such conditions, the most we might say is that the whole tangled web of inner resources constitutes a kind of miniature dynamical re-construction of the relevant aspects of extra-neural reality. But if we cannot go on to be more specific—if we cannot also assign finer-grained representational roles to component states or processes—then the explanatory value of the analysis is significantly (perhaps fatally) reduced. (Clark & Grush, 1999, p. 12).  

In this quote, Clark and Grush are ready to concede quite a lot to the antirepresentationalist. However, what they concede is not enough, or better said, is a (right?) move in the wrong direction. They are implicitly endorsing MR’s principle of reification. As we saw in section 1.1, this principle follows naturally from a realist reading of the connectivity principle. Clark and Grush are giving their backing to the idea that physical states of a system count as representational only when a parallel can be drawn between the states in question and the role they play in the establishment of a connection between the system’s input and output states. That is, they support the identification of specific physical states with the semantic roles they play. But note, nonetheless, that the acknowledgement of the possibility that the system may not be subject to de-composition is due in their view to the fact that the internal states of the system are massively and reciprocally causally connected. This however only represents a challenge that (representational) neurobiology would be willing to take on: Namely, the challenge of identifying and assigning representational roles to the sub-states and processes that neurobiology eventually discovers in the future. Or to put it another way, their acknowledgement simply tells against the accomplishments of today’s best scientific theories. The issue, in my view, cannot be simply reduced to an epistemological open empirical bet. It is rather an ontological dispute, such that the rules of the game can, at least in part, be settled on conceptually, whilst future neuroscience develops.

In section 1.1, I also introduced the principle of minimal lastingness as part of MR’s core instrumentality; a principle according to which, for a physical state to count as representational it must exhibit a degree of persistence that at least evens up with the degree shown by the system’s inputs to be represented. This must be the case since, for some aspect of the ever-changing world to be represented that may have only been presented to the system for a minimal period of time $t$, a representation of the world state must necessarily endure for a period longer than $t$. That is precisely what would allow the system to remain competent in

---

13 Their remark may be interpreted as pressuposing that representationality, after all, simply boils down to what counts as a good explanation. Although it may be argued that there cannot be further ontology in Clark and Grush’s position, I’ll press from another side. Thanks to Mark Bickhard for bringing this point to my attention.

14 In may be argued that some (minimal) representational states may be less transient than the external states to which they are tuned. In such cases, the principle of minimal lastingness may fail to be met
de-coupled, off-line modeling tasks, and be fleshed up beyond systems that are merely reactive, as Clark and Grush demand.

The question now is: Can cognitive activity possibly be accounted for in a transient manner? To answer this question we may endorse the following general stance. On the one hand, we may adopt a mechanistic level of explanation (e.g., Bechtel, 1998), where a part/whole (i.e., vertical) explanatory relation develops. And, on the other hand, we may opt for a dynamicist explanatory framework in terms of (horizontal) couplings; that is, couplings at the same level of physical aggregation.\(^{15}\) In my view, however, we may go for a dual route. As opposed to the vertical/horizontal dichotomy, we may call those systems that integrate simultaneously vertical and horizontal analyses, *diagonal systems*. In this way, diagonal modelling means that dynamicism can be applied at many levels of description simultaneously. However, there are no level-of-aggregation limitations intrinsic to modelling. We can have a dynamical model of quantum phenomena, infant’s reaching behaviour, or planetary motion. And crucially, we can also model systems diagonally such that sets of interdependent micro/macro quantities lawfully interact and can be studied. Thus, DFA models that integrate memory, perception and action (section 2) would constitute an example of simulated diagonal systems. A mechanical part/whole relationship can be applied, while microcouplings are honoured. That is, we can talk of self-organizing systems whose behaviour reduces to multicausal couplings, but where, crucially, nested timescales that correlate with different levels of aggregation govern the couplings themselves (see below).

In this way, “diagonally speaking”, talk of de-composition of the sort allegedly involved in a representationalist treatment may be accounted for mechanistically (cf. Clark and Grush, 1999, above), and relations between elements at different time scales may be kept track of. As we shall see below, such a dynamical time-scaled relation of constituency, where different levels of aggregation are entrenched, is at the heart of the analysis of the forthcoming notions of environmental reliability and primagensis. Put in the context of Clark and Grush’s line of reasoning, it is not the case, I shall try to show, that “the whole tangled web of inner resources constitutes a kind of miniature dynamical re-construction of the relevant aspects of extra-neural reality.” This is a problem that arises independently of the degree of complexity of the reciprocal causal influences found in the system. Let me elaborate.

Which are the theoretically relevant levels of analysis in the case of de-coupled activity? DFA may be endorsed by mathematicians simply because of the power of abstractions.

\(^{15}\) Hsi-wen Liu (2003), for example, considers this twofold choice.
However, I shall interpret DFA’s abstract variables and governing equations realistically.\(^{16}\) Taking for example the A-not-B error, the states that DFA posits are peaks of activation in activation fields. And it is the levels of activation in task input, specific input and memory input (see fig. 1 above) that are \textit{causally effective} in the production of the particular evolution profile in the working memory field. This will be pivotal when it comes to deciding, in order to answer the aforementioned question, what endogenous and exogenous levels of aggregation we’re to anchor at in the modelling of a diagonal system that is competent in the three different coffee-cup scenarios. If we grant that the use of activation fields is not merely heuristic, the question will ultimately be settled empirically. In this way, if working memory gets implemented, say, at the cell and molecular biology level, cognition may eventually be reducible to low level posits that get rezitized over low level variables such as, for example, neural firing rates.

At this point the possibility that we are confronted with a diagonal system where the exogenous dimension of the task is to be controlled at one level of aggregation, and the endogenous one at a different level, cannot be dispensed with. Therefore, the advantage of dynamically modelling a diagonal system is that instead of merely describing observable overt behaviour, we can actually try to get a fix on the underlying mechanisms that trigger the patterns of behaviour. The best open-minded approach then is to posit different sets of variables which may be low-level ones as in the case of neural patterns of firing, or high-level ones, as must be the case if we are to keep track of overt behavior. But before we have an actual model up-and-running we must set up constraints on which diverse levels of analysis are appropriate to study the phenomenon in question. That is, we must come to a decision as to whether, for example, the biochemistry of neurotransmitters needs to be honoured or whether modelling the firing rates of populations of, say, motor neurons suffices. How do we know that these lower level dynamics are essential to a full-fledged explanation of behavior. The litmus test that allows us to settle at a given level of aggregation is the existence of linkages between that level in question and the overt behaviour we aim to explain—e.g., the act of grabbing a coffee cup hindered by a pile of journal articles.\(^{17}\)

\(^{16}\) If we put it in terms of S&S’s coffee cup example (section 2), we are simply asking for a (non-linear) differential equation-governed model of the reaching data. However, it is not clear to me whether S&S endorse the nature hypothesis or simply the knowledge hypothesis (section 2). In accordance with the theory of minimal representationalism, I opt for a nature hypothesis reading, regardless of their intention.

\(^{17}\) It must be noted that overt behaviour in itself may also be the target of the same sort of decisions that apply, for example, to biochemistry and the firing rates of populations of neurons. In this way, we may ultimately reach a settlement at some level in the history of muscle movement, as opposed to the “level of” cup-reaching. For present purposes, I shall assume for argument’s sake the coarse level of description of overt behaviour as the target of scientific investigation. I am confident, however, that the reader will interpret this decision as inherent to theoretical and experimental design overall, rather than as a form of circularity. All that needs to be granted at this point is that Cognitive Science, writ large, does not rely on the autonomy of \textit{levels} (in Marr’s, 1982, sense), in such a way that representationalism cannot be vindicated \textit{a priori}, and remains an open empirical question.
As we saw earlier, DST defines stability as the *endurance* of a behavioral and/or neural condition under a given perturbation. The space delimited by the potential behavioral/neural conditions that the system may undergo is quantified by a set of (state) variables. The system is then said to *converge* to stable states from nearby points in phase space. Instabilities, on the other hand, were required for the system to evolve and reach new stable states. Neurally-generated activity was one of the reasons that could contribute to reaching stabilization. In this way, we may say that the benefit of bringing to view neuro-anatomical/chemical/physiological specification depends on whether constraints at those different levels find a correspondence with stable patterns of behaviour. Were such a correspondence to be discovered, a substitution of higher level parameters for lower level ones would be justified, and would add in an advantage in terms of specificity and explanatory understanding.

Bickle (2003) has recently presented an illustration of a successful linkage between lower level molecular mechanisms and stable overt behaviour. He forcefully argues that an explanation of “quantitative behavioral data at the level of biochemical pathways and intracellular molecular mechanisms” is already on offer. In his book, he displays a comprehensive body of evidence that shows how particular behavioural data can be explained at the level of the molecular mechanisms that implement long-term potentiation (LTP). Stable patterns of behaviour, in this case, are subject to the scrutiny of endogenous molecular activity. In this way, instabilities and subsequent stabilization processes, such as the patterns of stability that mediate the shift that takes place from short-term to long-term memory—i.e., cognitive psychology’s “memory consolidation switch”—can be successfully explained. The moral is that results of cellular and molecular manipulations in control experiments must be included when inserting neurobiologically plausible parameters in our DST models.

All this drives us to the following considerations. There are certainly many interesting microscopic levels of analysis that will ultimately be explained—e.g., micromechanisms involved in limb position, proprioception, etc., etc. However, all that is required is that the level of description that we choose to settle on is empirically shown to be linked to overt behaviour in such a way that stability at the level of the phenomenon under study can be appropriately accounted for. In this way, we can endorse dynamicism, and decompose a diagonal system at different interlinked levels. Thus, if we discover, let’s say, that Na+–K+ pumps’ opening related activity fails to be positively correlated with the stability of the pattern of behaviour under study, we may conclude on these grounds that that particular biochemical mechanism or process is irrelevant. 18 We may likewise conclude that biophysical constraints related to limb

---

18 Notice however that this doesn’t mean that the level of biochemistry that underlies neural activity in itself is not the appropriate level of description for DST modeling. Other mechanisms/processes at the same level of aggregation may well turn out to play a critical role.
weighting constitute or fail to constitute a correct description to be included in a multi-causal dynamical equation.⁰¹

In short, the existence of a linkage between a level of aggregation and stable behaviour is what will determine the validity of the level in question. In this way, if we are to bridge the gap by means of DFA principles between different levels of aggregation in a diagonal system, how different brain structures relate to the global (off-line) behaviour of the system will need to be considered. As molecular neurobiology progresses and parameter values are ecologically pinned down, finer-grained neurophysiological processes will putatively be reflected in sets of differential equations. Relationships between these and coarser-grained modelled mechanisms may then be studied.

A clear example is “single cell” neurophysiology, where some astonishing progress has been made in accounting for phenomena that traditionally were thought to be only explainable by cognitive neuroscience.⁰² By microstimulating clusters of neurons with tungsten stimulating electrodes, electrical activity can be induced in very small numbers of greatly specialized neurons, and the real-time behaviour of animals can be studied. Methodologically speaking, one of the long-range objectives of standard neurophysiological techniques such as microstimulation is to “create realistic experiences and mental operations artificially, by directly activating known circuits of neurons in the brain in the absence of the external inputs that normally elicit such mental operations” (Liu and Newsome, 2000, R598; quoted from Bickle, 2003, p. 164; emphasis added). In this way, one of the purposes of microstimulation experiments is to produce dissociations that can tell apart endogenous as opposed to exogenous related brain activity. In a series of experiments, Newsome and colleagues performed dissociations of this sort on primates successfully inducing stimuli qualities on the basis of neuronal microstimulations themselves, rather than exogenously in terms of the tangible features of environmental stimuli. This sort of neuroscientific studies aims at showing that endogenous activity, rather than external exogenous activity, is what really matters to understand cognitive activity. Although there is a significant degree of correlation between information provided by the environment and endogenously generated activity, crucially, the two can be dissociated:

A plausible hypothesis is that microstimulation evokes a subjective sensation of motion like that experienced during the motion aftereffect, or waterfall illusion. … Motion therefore appears to be a quality that can be computed independently within the brain and “assigned” to patterned objects in the environment” (Salzman et al., 1992, p. 2352; quoted from Bickle, 2003, p. 198).

---

⁰¹ Equally, it must be noted that a macroscopic approach to the motor behaviour of a diagonal system must not necessarily slot in the actual kinematics of the grabbing action. Nor is honouring anatomical details beyond joint configurations or centres of gravity a priori demanded.

⁰² See Bickle (2003) for a comprehensive review of core results. The next paragraph distils key aspects of chapter 4 of his book. The reader is urged to visit section 5 of that chapter for the minutiae.
Bickle interprets Salzman et al.’s microstimulated dissociations as follows:

They are suggesting that visual motion qualia are generated in the brain and attached to internal visual representations of external objects. Happily, in ordinary circumstances, our “internal assignments” of features to representations tend to correlate with features and relations of the objects represented. Natural selection was crueler to creatures whose “internal assignments” were more haphazard or skewed. But under appropriate conditions, our internally generated and assigned qualia and the external environmental features can be dissociated” (Bickle, 2003, p. 198; emphasis in original).

3.2 Reliability of environment and primagenesis

These two quotes implicitly furnish us with the basic toolkit to introduce the key concepts of environmental reliability and primagenesis. Notice that the induction of behavioural experiences is directly applicable to our discussion insofar as microstimulations at the cellular and/or molecular level have already been shown to be linked to overt stable behaviour. Let me start in inverse order, and pave the way for primagenesis first.

In order to fully appraise the bearing of the notion of primagenesis upon the representationalism/computationalism debate it will be useful to distinguish it from other closely related suggestions in the neuroscientific literature. Llinás and Churchland (1996) propose the concept of endogenesis as an account of the broad initiative that is at the heart of dissociation studies. Endogenesis stresses the fact that cognitive activity is the result of endogenous neural activity: “sensory experience is not created by incoming signals from the world but by intrinsic, continuing processes of the brain” (Llinás and Churchland 1996, p. x; quoted from Bickle, 2003, p. 198; emphasis added). Likewise, what I shall dub “primagenesis” (primary endogenesis) stems from the notion of dissociation. Although related to endogenesis, what really matters in the case of primagenesis is continuity itself, rather than the fact that the signals that induce sensory experience come from sources internal to the system. Thus, I shall define primagenesis as follows:

Let level A be the lowest level of aggregation of a system that allows for dissociation empirically and that our best neuroscience can successfully integrate with DFA principles. Primagenesis then refers to the continuous set of processes at level A that remains linked to higher-level performance and produces a given piece of stable behavioural data.

It should first of all be noticed that primagenesis significantly establishes a crossing between continuity and stability. Nevertheless, the differences between Llinás and

---

21 Bickle exploits dissociations in the context of Newsome’s microstimulation-motion studies with monkeys in order to show the induction of qualia from visual motion. In his sense, endogenous activity is essential to understanding qualitative sensory experience. For present purposes, nothing hinges on phenomenal experience issues. My interest relies on dissociation studies and the external/internal continuum.
Churchland’s concept of endogenesis and my concept of primagenesis are twofold. On the one hand, primagenesis, although putatively found in complex organisms strongly linked to endogenous cortical activity, should not be seen as explicitly stressing intrinsic brain processes against exogenous activity. The emphasis is not on endogenesis, but on being primary, in the sense of targeting whatever level of aggregation is the lowest one where a linkage along Bickle’s aforementioned lines is obtained. In this way, were we to observe an external stimulus directly linked to the internal state thought responsible for it, it would count as a genuine case of primagenesis. What really matters is the fact that the patterns of activation that evolve in cortical regions are to be interpreted in terms of neural signals in a continuous fashion. If we follow this suggestion, whether those time steps take us to activations further down within the cognitive system as opposed to the periphery of the system, and ultimately to exogenous activity, becomes a side issue. Likewise, it is noteworthy that primagenesis also differs from a straightforward arithmetic sum of different decentralized contributions. Although dynamicism stresses the fact that external non-cognitive factors and internal ones are to be treated on a par, by exploiting the insights of microstimulated dissociations, primagenesis makes the most of the state the system is in at (immediately) prior time steps; specifically, of the state of the system at the closest causally efficacious and continuous level A of processing. Thus, rather than claiming that endogenous and exogenous inputs to the system interact non-linearly and their contributions are to be summed up in the cognitive equation, we may say that level A reflects the activation generated by those contributions at immediately previous time steps. Since, in principle, different time scales can interact in a diagonal system, the lower the level of aggregation we empirically identify with level A, the temporarily closer level A will be to triggering the pertinent effect/s.

On the other hand, a second reason to differentiate primagenesis from endogenesis resides in the fact that endogenesis may run the risk of ignoring the decentralized (embedded, embodied) character of cognition, essential to the dynamical modelling approach. Llinás and Churchland claim: “Incoming signals from sensory receptors keyed to external physical parameters serve only to “trellis, shape, and otherwise sculpt the intrinsic activity to yield a survival-facilitating, me-in-the-world representational scheme” (ibid.)” (Bickle, 2003, p. 198; emphasis added). In this way, endogenesis plays down the role of external contributions, regarding them as little more than an act of rough-hewing. By contrast, primagenesis understands the external dimension as part of a non-trivial multi-causal spread chain (see Calvo Garzón, and the references therein). This is not inconsequential since note that endogenesis, as

---

22 A different matter is that because higher-level, de-coupled, activity requires the recoding of information at multiple levels, ultimately it will be effectively endogenously generated. On the other hand, some minimal representationalist may well argue that peripherically transduced information in relatively simple agents may lead to forms of encoding that enter in direct (representational) correspondence with external features of the world. But since our target is de-coupled activity I shall leave that consideration aside.
introduced above, still is minimally representationalist, whereas, as I shall argue next, primagenesis need not be.

Primagenesis would find a natural implementation in dynamic field terms. Notice that it is crucial for DFA’s explanation of the infants’ reaching behaviour (section 2) that some inputs to the system come from sources which are not external to the field. That way, the field is allowed to self-organize by neighbouring patterns of excitation and lateral inhibition. Primagenesis, in this way, would explain why peaks within the activation field happen at times to self-sustain; a crucial feature if we are to explain older infants’ non-perseverative behaviour. Notice that in the case of 12-month-old infants, who don’t make the error anymore, although the input values coincide with the top panel ones of figure 1 (section 2.2), evolution in the working memory field differs. As the object is hidden, there is no decay of activation at B, remembering thus the act of hiding the object at the B location. This is due to the different way in which activation excites/inhibits neighbouring/further away sites. In the case of older infants, endogenous interactivity reverberates by means of inhibition/excitation connections, making particular inputs to the system have less weight with regard to the overall output behaviour. My working hypothesis then is that a primagenetic level A of processing that relates to the memory fields may be credited in the differential equations that explain perseverative as well as non-perseverative behaviour.

We shall go back to this issue in greater detail shortly, but for the time being, a closer look to Thelen et al.’s motor field approach will shed further light upon our second key concept: Environmental reliability. As stated by DFA, the A-not-B error is due to a motor planning process in the context of a perception-action dynamical loop. According to Thelen et al., the system is endowed with a history which is to be understood continuously: Motor planning is continuous and the planning evolves under the continuous sway of perception. Now, claiming that the system has history may run the risk of turning the model into an information-processing architecture, where historical information is manipulated in algebraic and/or statistical ways. However, under the lens of primagenesis, we may also say that history reduces to previous (level A continuous) time steps, in such a way that what we read as “having history” simply boils down to exploiting a reliable environment. In this context, what I shall dub “environmental reliability” (ER) simply refers to the uncontroversial fact that features of the external environment tend to correlate in appropriate ways. Or, putting it slightly differently, the world presents regularities that allow us to cut across idiosyncrasies, and crucially these regularities are presented reliably across diverse exposures. Rather than reflecting a correspondence with environmental factors, the model points to the lack of a tuning to underlying environmental regularities. In this way, we are able to fine-tune to the world’s regularities, not because we discover an underlying structure, but rather because it is basically reliable. We may then say that an environment is reliable if there’s a matching between external
stuff and endogenous activity. But the locus is never the external world, but simply the activity
that obtains at level A of processing.

The approach put forward here is partly Gibsonian in spirit, but the reasons may not be
fully appreciated at once. ER clearly sets apart the ingredients proposed herein from a situated
cognition fabric. It is not the fact that most of the information is out there what really matters,
but rather the fact that the environment happens to be reliable, even though external regularities
may not be actually accessible when acting. Situated cognition exploits the fact that cognitive
activity cannot be detached from the surrounding environmental factors. That is what gives the
theory its direct realism flavour. Representational economy is due to the fact that a
considerable amount of data happens to be accessible in the vicinity. In this way, there’s no
need to keep a large stock of enduring representational elements that would allegedly stand for
what is already present in a relatively stable form. However, my position is not committed to
direct realism (nor to indirect realism, for that matter). Notice that direct realism could also be
interpreted in representational terms. We can simply interpret in contentful terms the
information of the system’s “input layer”, independently of whether there’s a “hidden
representational layer” or not. If we are to remain genuinely antirepresentationalists, the issue
must run deeper. According to the approach advocated here, thinking of those aspects of the
world that remain intact as the reason why no enduring internal states are needed is the wrong
way to look at things. What really happens, the story would go, is that cognition works without
enduring states not because some external aspects remain intact, but rather because those that
do happen to remain intact, as well as those that do not, are all reliable. The insight from the
notions of primagenesis and ER is then that “off-line” predictive power doesn’t arise because
experience is stored prior to motor action. But, nevertheless, neither is a representational
cosmic fluke involved: “Happily, in ordinary circumstances, our “internal assignments” of
features to representations tend to correlate with features and relations of the objects
represented.” (Bickle, 2003, p. 198). Put bluntly, situated cognition exploits the fact that a
sizeable amount of environmental data remains stable beyond idiosyncrasies, whereas
antirepresentational dynamicism exploits the fact that internal system states change
continuously in relation to changes at a primagenetic level A under a reliable environment.

3.3 Towards a general theory of antirepresentationalism
We are now in position to take action in response to the difficulty raised by S&S (section 2). In
a nutshell, and to unveil the structure of my gambit, the crucial step to take in is that once we

23 It is actually the fact that in many tasks the relevant environmental factors are usually out there what
fools us, and doesn’t let us realize that the external object has nothing to do, as Llináš and Churchland’s
concept of endogenesis already hinted at!
grant primagesis, CD, the thesis of content discontinuity (section 2.2), dissolves. To remind the reader, S&S claimed in the context of the three coffee cups example:

Does the motor approach capture everything about this situation? The answer is “no”. This approach fails to capture differences in the representational states underlying these movements. In particular, in one case, there was a high degree of certainty regarding where to move, in the second case, the decision of where to move was much less certain, and in the third case, the decision took on a new flavor— rather than being generated based on visible information, the decision was generated based on a longer-term memory of the cup’s location. (S&S, 2003, p. 398)

As we saw in section 2, the basic working hypothesis in S&S’s approach could be stated as follows: Whatever model ultimately succeeds in accounting in dynamic terms for, say, the coffee cups example, will have to incorporate the inherently discontinuous character of representational states. I argued that in accordance with CD, for a cognitive system to count as representational it was sufficient, although not necessary, that activation states could be discontinuous. Two states were discontinuous when field evolution did not require the activation of all locations in between the states to build up. Unfortunately for the sympathiser of representations, someone’s *ponens* is tantamount to somebody else’s *tollens*. Read “discontinuity” versus “continuity”, and, put bluntly, the problem with S&S’s explanation of the third coffee cup scenario is that it fails to meet the “(cognitive) continuity hypothesis,” (CCH; section 1, above). CCH incorporated a homogeneity constraint according to which (i) higher-level cognition was on a par with bodily and neural processes, and could thus be accounted for in terms of continuity, temporality and non-linearity; and an interactivity constraint that says that (ii) multiple level interactions take place in a continuous fashion all the way down from high-level cognitive activity to perception, motor control and non-cognitive processes in general. Both homogeneity and interactivity made continuity one of the pivotal concepts in explaining behavior and cognition. However, although S&S respect the interactivity hypothesis, they mistreat the homogeneity constraint.

Paradoxically, nonetheless, S&S were careful enough to claim that endorsing CD did not seem to go against CCH:

This is not to say that the processes underlying these two representational states were not continuous, nor does it imply that continuous change of representational state never occurs (e.g., continuous state change does occur in many mental rotation tasks). The central point is that the content of representational states does not necessarily take on this continuous character. This stands in contrast to the motor system which must always evolve continuously in “content”. Thus, a pointing response to the “left” must necessarily evolve through intervening locations en route to the “right”. (*ibid.*, p. 399)
Towards a general theory of antirepresentationalism

In section 2, a distinction was established between a concrete dynamical system, a real object that changes over time in the physical world, and the abstract dynamical system that may be instantiated, that is, the mathematical model that describes formally the actual changes undergone by the concrete system. In order to meet CCH, a DFA abstract model should describe those continuous changes undergone by the concrete dynamical system which are at the heart of the homogeneity constraint. Once we grant that the “memory consolidation switch” reduces to cellular and molecular mechanisms (Bickle, 2003), the description of the DFA model will be correct only if it respects those continuous aspects of neurobiological changes that are empirically linked to the reaching behaviour. The interactivity constraint, on the other hand, once level A is empirically settled upon, won’t suppose the inclusion of discontinuities, since a correct instantiation of the abstract dynamic system will reflect the continuous processes at level A that remain linked to higher-level behaviour and produce the actual piece of stable reaching behaviour.

On the other hand, Erlhagen and Schöner (2002) mentioned as a limitation of DFA the fact that particular movement parameters are discrete (despite the fact that the concept of a dynamic field stresses continuity in movement parameter space). For example, choosing to move the right or the left arm is interpreted as a discrete selection behavior. In my view, that is what has prevented many from seeing the solution to the puzzle. Namely, the fact that we try to simulate the cognitive process at such a high level of description. A lower level ecological setting would show that activation values in movement space are not an all-or-nothing choice. This means that discontinuous change of the sort S&S require may amount to an incorrect or less accurate description of the phenomenon. When S&S observe that a discontinuous shift can mirror the change that takes place when a target disappears at one location and appears at a different one, and claim that “this can occur even though activation never builds up at intermediate reachable locations”, they may be aiming at the wrong target. Were simulations to target a primagenic non-discrete level, these worries would not arise. This is specially evident in S&S’s third case, where although no visible information is available, the decision to reach the hindered cup is generated based on a longer-term memory of its position.

Now, under the light of primagensesis we can see that a diagonal system that meets CCH’s homogeneity and interactivity constraints is actually orthogonal to S&S’s demands. Contra S&S, CD as a matter of fact, I claim, does go against CCH. Level A primagensesis cancels out S&S’s acknowledgment that implementing brain structures are continuous, since once we assume that level A processes are linked directly to overt behaviour, CD dissolves. States of a diagonal system subject to a level A analysis do not necessarily have the discontinuous

---

24 It must be said that this in fact is explicitly pointed out by Erlhagen and Schöner (2002).
character putatively required. On the contrary, the key role of primagenesis is to approximate motor continuity and higher-level alleged discontinuities.25

On the other hand, enduringness is usually interpreted as the coding of discrete sensory stimulations. A familiar illustrative example would be the coding of linguistic spatial or temporal relationships, where enduring discrete constituents are exploited so that systematicity and productivity can emerge via combinatorial means. However, it should be studied in the context of primagenesis.26 Not only does dynamicism call into question discreteness, by its emphasis on continuity. Primagenesis also calls into question enduringness, as stated under the principle of minimal lastingness.27

Crucially, discreteness and discontinuities at the higher level do not exclude continuous neural behaviour. Note that in Thelen et al.’s rendering of the A-not-B error, in order to correctly describe perseverations in reaching, noise and thresholds are included. These are simplifying assumptions that arise because of ignoring finer-grained neural detail in their modelling. Were thresholds not to be included, continuities at the neural level would need to be appropriately inserted in their mathematical equations. In Thelen et al.’s A-not-B simulations, the activation field, which is meant to match the spatial vicinity of the infant, is assumed to be in a resting state, with all activation values equal to a parameter $h$. It is assumed for simplicity’s sake that the infant’s reaching is produced whenever a pre-specified threshold is exceeded. 12-month-old infants non-perseverative behaviour is then modelled by making appropriate changes to $h$. If an activation value goes over the threshold, the infant reaches towards the location (A or B) that corresponds to that activation site. The parameter operates by modulating the degree of cooperativity of the different sources of information available in the task (see Thelen & Whitmyer, 2005, for the details), and can drive subsequent performance in two ways. If $h$ is not high enough, the activation peak decays relatively fast. This means that the system is unable to self-maintain its behaviour when the external input is no longer present. However, once $h$ exceeds a certain parameter value, the activation field manages to reach a state of self-maintenance, such that the activation peak remains self-sustained even though the external input is no longer present. This second sort of state behaves as a sort of memory device, and the activation field is said to be in a “cooperative regime”.

25 Note that this does not mean that a level of description that avoids the exploitation of discontinuities automatically qualifies as explanatorily appropriate. This will depend on having the right sort of links with behaviour as was suggested earlier.

26 In Calvo Garzón (2004b) I articulate a notion of constituency (orthogonal both to classical and connectionist approaches) where combinatorial behaviour gets grounded in sensori-motor activity and the parameter of time.

27 An anonymous referee wonders whether by questioning discreteness my position also implies a departure from recent strong reductionist programs, such as Bickle’s (2003). Although that’s a possibility that cannot be discarded beforehand, an appropriate answer would require examination on a case-by-case basis. Notice that even within an agreed-upon level of description, some mechanisms/processes will be discrete and others continuous.
Parameter \( h \), unfortunately, does not have a psychological counterpart. Realistic parameter values, nonetheless, can be assigned, once we settle back, according to primagenesis, at a level A of description. Continuous dynamics of the memory field, related both to the decay of activation that takes place in the non-cooperative regime, and the growth of activation that drives the system in a cooperative regime towards self-sustaining states, will in principle be mapable onto ecological factors, once the real-time interactions among LTP molecular mechanisms is understood in detail.\(^{28}\) In short, parameter \( h \) is not to be interpreted psychologically, but rather it is to be recast continuously at level A of processing. Again, the positing of discontinuous states is due to the failure to note that the choice between an A or a B position relies on a continuum where level A system states for A or B are not discrete choices but two points of a continuous set of parameter values.

The idea that representations are still involved in the A-not-B error is due to the fact that motor memory, which is crucial for a dynamical explanation of the phenomenon, is modelled at a level of aggregation that does not allow us to exploit primagenesis. S&S seem to be shooting themselves in their feet when they claim:

[An] appropriate level of description must be closely linked to the stability of the behavioral pattern under study. This follows from the goal of a dynamic systems model – to capture how the stability of behavioral states changes over time. Thus, if we find, for instance, that membrane mechanisms work the same way irrespective of whether a particular behavioral pattern is currently stable or not, then we are not looking at a level of description specific to the phenomenon that we aim to understand. (S&S, 2003, p. 398; emphasis added—see fn. 18, above.)

The general framework advocated here involves a radical alteration of how we understand cognition. Rather than focusing on the simulation of the abstract de-coupled processes allegedly involved when performing a given cognitive task, the emphasis should rely on how the values of primagenetic (level A) magnitudes of a diagonal system vary in time.

The problem with S&S’s line of reasoning stems from the fact that their reading of the motor approach fails to establish a solid connection with a level of aggregation equivalent to primagenesis’ level A. Why, in short, should we keep constant the motor variables in the dynamic equation? The criterion I propose would be: Keep constant whatever part of the overall system under study (brain, body, environment), call it part X, such that the maximum amount of data can be dealt with in terms of continuities while remaining behaviorally adequate. If we hit empirically level A continuities and we acknowledge that the environment just happens to be reliable, then what we obtain as an extra bonus is a dismissal of content discontinuity. In short, claiming that de-coupled, off-line dynamics are not necessarily continuous is due to having an incorrect level of description/aggregation. The fact that level A

\(^{28}\) See, for example, Miller, Erickson & Desimone (1996) for a neurophysiological rendering of the maintenance of prefrontal activation after external stimulation vanishment.
Towards a general theory of antirepresentationalism

is continuous entails that the internal state is not about an external missing object, but rather it simply codes dynamically at time $t$ the state of the system at a prior time step.

Finally, it is also worth pointing out that primagenesis need not coincide with first order correlational structures. They may actually overlap in case level A happens to tune reactively to first order correlations in the data pool. But notice that this will depend on the particular level of aggregation level A gets anchored at. Independently of whether correlations are first-order or any higher n-order, the key test is to know whether the link with the behavior we aim to explain is preserved or not. Otherwise we would have a level of aggregation that effectively relates to low level correlations but where no explanatory power is obtained. The relevance of this point may be easily overlooked. The pertinent reading of the CCH hypothesis is that cognitive activity is realized in systems subject to coupled dynamics at the relevant level of causal organization that primagenesis indicates, whatever that level comes to be.

Now, once the whole landscape is reconsidered under the light of the above considerations, we can see that the very first question we asked when framing the classic Piagetian A-not-B scenario is ill-posed. That is, we were phrasing cognitively the following question: What is it that infants represent of the world that allows them not to perseverate in their erroneous reaching? This rendering of the situation presupposes an answer in which the dynamic field of activation obtains stable representational states, of the sort that can minimally represent the environmental state being asked for. On the contrary, my proposal simply involves a continuous transformation along a sensory-to-neural pathway, such that MR principles fail to be met. Thus, the question appropriately recast would be: What sensory-to-neural continuous transformations permit infants not to perseverate in their erroneous reaching?

In my view, the burden of proof is always placed on the sceptic’s side and discussion proceeds without further ado. By contrast I reject that burden, and put matters the other way round: There are no representations because they require discontinuities, and level A dynamics are always continuous. In order to make conceptual progress we need a principled distinction which gets couched in terms of degree of continuity. The less discontinuous, the closer we are to nonrepresentational dynamics. Therefore, representations and computations deflate as we approach primagenesis, and they finally evaporate. Information-processing representationalism, I conclude, is not a paying simplification of antirepresentational dynamicism.

4. CONCLUSION

The lesson to be extracted from Piaget is that higher level cognitive activity grows out of a sensorimotor base. Unfortunately, this acknowledgement stopped right there, at the honouring of its provenance, in such a way that as cognitive tasks grow in complexity, subsymbolic and/or symbolic resources were taken to be unavoidable posits. The lesson to be extracted from an antirepresentationalist reading of dynamic system theory is that coupled dynamic processes
remain equally grounded at any age. There is no stage in development where cognition ceases to be embedded and embodied.

The spirit of this work has been to advocate a new view of dynamic systems theory that explicitly reckons antirepresentationalism as probe bed for cognitive science research and philosophical inquiry. After the disavowal of (Skinnerian) behaviourism, Cognitivism, the dominant paradigm nowadays, is characterized by the *sine qua non* condition that internal representational states and computational processes, in one form or another, are assigned an explanatory role. In this article I have defended the thesis that physical diagonal systems, as specified by sets of differential equations, are temporarily and spatially continuous, such that an indefinitely large set of information can be processed. The result is a dismissal both of computational (effective) descriptions of the processes involved, and of a *minimal* representationalist treatment of the system states.\(^{29}\)

Nevertheless, the main objective of this article is not to argue that dynamic system theory is a new paradigm, but rather to manifest what it still lacks or what it should look like to become a genuine new paradigm. The discussion of semantic content in section 3 shouldn’t be interpreted as final arguments against the theory of mental representations, but rather as pilot steps towards a general theory of antirepresentationalism. (Un)fortunately, many vexing problems remain to be addressed (see Calvo Garzón, submitted). Readers are thus requested to do their primary reading in the context of the potential paradigm shift that researchers in the dynamicist movement have been calling for during the last decade.

On the other hand, it doesn’t escape me that the line of research favoured in this paper has indirectly been attacked as reminiscent of logical positivism and strong behaviorism (e.g., Pelphrey and Reznick, 2001). At this point, we can simply say that only time will tell if neuroscience needs to retain psychological posits or not (but no more than new research hypotheses and testable predictions will tell). It is my hope that the continuous and situated approach of dynamicism may lead us back to the main road by eschewing, rather than revising, the (computationalist) function-approximator approach. Classically posited questions, for example, such as to whether “complex cognitive functions … depend on a mixture of statistical and algebraic (rule) mechanisms” (Marcus and Berent, 2003, p. 53) may simply vanish once the continuous interplay that dynamic system theory calls for between brain, body and environment is modelled *in sufficient detail*. The proof, however, is in the pudding, and the payoff of the notions of primagenesis and environmental reliability is the fruit they bear for cognitive science, both empirically and conceptually. How much further we can go from the foundations

---

\(^{29}\) The rejection of *minimal* representationalism is what prevents this primagenesis-endowed dynamicism from collapsing into reductionism, however radical it may be. If the arguments put forward here are correct, a further-reaching thesis holds, namely, the thesis that cognitive science is deeply mistaken about the true nature of cognitive activity. Cognitive states, as described by a minimal theory of mind/brain representation, simply don’t exist.
provided by antirepresentational dynamicism remains to be seen. The research reported here is an attempt to induce such a shift of paradigm in cognitive science research agendas.

References


Calvo Garzón, F. (submitted)

Calvo Garzón, F. The Collapse of Scientific Intentional Realism.


Figure 2. Dynamic Field Theory of the A-not-B error. Top panel shows time-dependent changes in a spatial working memory field (d) in the context of three inputs (a, b, c) during the first B trial for an 8-10 month-old infant. Bottom panel shows changes in spatial working memory (h) in the context of the three inputs (e, f, g) during the first B trial for older infants. (From S&S, 2003, p. 401. See ibid. for details.)