avant

2/2014

trends in interdisciplinary studies

T R E N D S I N I N T E R D I S C I P L I N A R Y S T U D I E S

AVANT 2/2014

Enactivism: Arguments & Applications

EDITORS
Tomasz Komendziński, Przemysław Nowakowski, Witold Wachowski

TORUŃ–WARSAW

ISSN: 2082-6710

AVANT. The Journal of the Philosophical-Interdisciplinary Vanguard
Trends in Interdisciplinary Studies
Enactivism: Arguments & Applications. Vol. V, No. 2/2014

This online version of the journal is a referential version.

The cover design by Karolina Pluta. The picture on the cover: "Garden Rose" by Robert Lemay. The graphics inside by Antoni Domański.

Address of the Editorial Office:
skr. poczt. nr 34, U.P. Toruń 2. Filia,
ul. Mazowiecka 63/65, 87-100 Toruń, Poland
www.avant.edu.pl/en *avant@avant.edu.pl*

Publisher: Centre for Philosophical Research
ul. Stawki 3/20, 00-193, Warsaw, Poland *www.obf.edu.pl*

Academic cooperation: university workers and PhD students
of University of Warsaw and Nicolaus Copernicus University.

The Journal has been registered in District Court in Warsaw,
under number: PR 17724.

AVANT is listed in: SCOPUS (Elsevier), The Directory of Open Access Journals (DOAJ), Index Copernicus Journals Master List, Central and Eastern European Online Library (CEEOL), The Central European Journal Of Social Sciences And Humanities (CEJSH), EBSCO Publishing, PhilPapers, BazHum: Virtual Library of Science, Arianta, and others.

Table of contents

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en

Preface

Dear Readers,

we are pleased to present this year's second issue of *Avant*, entitled "Enactivism: Arguments & Applications".

The main section is devoted to the dispute over what the essence of enactivism is, and what role it should play in research on cognition, as well as the applied side of enactivism. We present papers by Kenneth Aizawa, Pierre Steiner, Krystyna Bielecka, Fred Cummins, Qing Li, Ian Winchester, David A Reid, Ralph Ellis, Robert Briscoe, Jean-Luc Petit, Shaun Gallagher and Matthew Bower.

The issue also includes three interviews: with Shaun Gallagher, Robert D Rupert, and Robert Lemay, as well as book reviews by Patricia Grosse and Christopher Drain.

Editorial Board

Torun–Warsaw, Autumn 2014

introduction

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0001

Cognition as shaking hands with the world. Introduction

Przemysław Nowakowski* & **Tomasz Komendziński**[1]
Nicolaus Copernicus University
**pnowakowski[]avant.edu.pl*
Received and accepted September 2014; published Autumn 2014.

Your fingers would remember their old strength better…
if they grasped your sword.
Gandalf
(from the movie *The Two Towers,*
second part of *Lord of the Rings* movie*)*

One of the most common questions in today's cognitive studies is the one regarding embodied cognition. The answer to this question draws our attention to many factors, including bodily actions, which also work to *embody* cognition. With this in mind, enactivism is *included* in discussions of embodiment. In the current issue we present texts in which a focus on enactivism itself is the leading topic.

If one were to describe declaratively the latest trend in cognitive studies, one would frequently refer to it as "embodied cognitive science"—sometimes with the addition of "radical embodied cognitive science"—or "enactive cognitive science". However, attempts at answering the question regarding the relations of range and meaning between these terms set in motion a never-ending discussion. The issues connected with embodied cognition and enactivism tend to refer back to areas outside the field of cognitive studies. Including the category of embodiment (often quite contingently connected with situated and distributed cognition) within the context of enactivism, whose methodological and historical-ideological status is ambiguous (is it a methodological approach? a trend?), generates additional problems and questions.

An attempt at introducing some order into the situation would require setting clear criteria and conducting a detailed notion analysis. One should always take into account various ways and contexts of using the categories of embodiment and enaction, which seem to lead us into the even broader waters of interdisciplinary studies. This is compounded by the additional problems

[1] Transl. Ewa Bodal.

faced by the fields in which the aforementioned discussions are conducted—including cognitive studies and interdisciplinary studies, and especially philosophy and psychology in particular. However, the present introduction does not aim at bringing order to these various levels on which enactive issues appear. Furthermore, there is also no guarantee that ordering is reasonable, necessary, or efficient for research, and whether this is even possible.

A certain common notional basis for enactivism is often pointed towards; it is comprised of such notions as autonomy, sense-making, structural coupling, self-organisation, agency, action, and sensorimotor dependencies. It is still a long way from showing the relations between these notions and from unifying enactivisms into one common theoretical proposal in a satisfactory manner. However, not only does this not stop the researchers, but it also encourages them to further—especially critical—studies, which will allow enactivism to discover itself anew.

At present, enactivism is explored in several partially different directions. From the radical biological roots of the theory of autopoiesis (Varela and Maturana), currently frequently referred to in the context of research on artificial life (Di Paolo, Froese), to research on cognition carried out through the sensorimotor system (Noë, O'Regan), to the role of interactions in social cognition and sense-making (de Jaeger, Gallagher), to seeking models of mind and the role of procedural knowledge in cognition (Hutto), a strong emphasis on the constructive character of cognition (Maturana) or pointing towards the key role of self-organisation and emotions in cognition (Ellis, Newton). The hereinabove distinguished notion basis refers precisely to all these issues. These remarks do not solve the issue of the variety of enactivism, but only point towards it.

"Enactivism" has its terminological source which is turning attention towards the role of very broadly understood actions as the key to understanding what cognition is. Shaun Gallagher (2013: 209) writes:

> The enactive view of human cognition starts with the idea that we are action oriented. Our ability to make sense of the world comes from an active and pragmatic engagement with the world, along with our capacities to interact with other people.

McGann et al. (2013) are comparing cognition to a handshake and to dancing. They write that we have to use cognition when it is taking place, when it constitutes the action we are currently performing. This differentiates the enactivism from the concept of action. Much of research on action is—at least in light of some enactivist works—anti- or at least non-enactivist. Enactivism equates cognition with action, but it defines the criteria of "action" in its own way, focusing on its very performance. Other concepts focus on the mechanisms that make action possible.

The enactivist, equating cognition with performing actions, finds the results in research on cognition being something akin to catching a flying ball. To put it differently—analogically from Gandalf's words to King Théoden—cognition happens when, in a specific context and with a specific tool (here: a sword), we start performing an action. This framing has both its advantages (e.g. observing cognition *in statu nascendi* may reveal many properties of the act of cognition), and disadvantages (it is possible that many of the processes that make acting possible do not reveal their basic properties only through action). Although the enactivist proposal seems very inspiring, is it enough for a revolution in, or a unification of, cognitive studies?

What seems to connect the enactivists, besides pointing towards action or towards a certain particular practice of cognition, is rejecting the existence of mental representations or their necessity for explaining the essence of cognition. However, the situation is more complex than it may initially seem. On the one hand, not all enactivists reject the concept of representations (e.g. Natika Newton), and on the other hand—a certain issue is located in the very status of representations criticised by enactivists, as well as the potential concepts of representation insusceptible to the critique of the enactivists (see: Steiner in the current issue). The same pertains to the notions we have listed above. Not every enactivist refers to such notions as self-organisation, autonomy, or sense-making, and the authors who do reach for them differ in their interpretations of these notions.

The present issue consists mainly of ten articles that are ten standpoints on enactivism: not only from the point of view of the critics (Aizawa, Steiner, Cummins, Bielecka) and the proponents (Gallagher & Brower, Ellis, Li and Winchester), but also spokespeople for the moderate approach (Reid, Briscoe, Petit). Let us briefly outline the contents of the main part of the volume.

Aizawa (in this issue, as in: Aizawa 2014) presents doubts regarding enactivism, focusing on the category of "cognition", central to cognitive studies. Having shown why differentiating between cognition and behaviour is important for cognitive scientists, Aizawa analyses a number of works by enactivists in order to show that they frequently mistake cognition for behaviour. Steiner considers the enactivist critique of representationalism, showing that on the one hand, enactivists do not present a unified concept of representation in their critique, and, on the other hand, there are possibly such (non-referential) concepts of representation that are immune to this critique. Bielecka proves that the radically externalist theory of content, as present in Manzotti's research, is not possible to support, pointing towards the doubtful concept of hallucinations by this author. Cummins criticises the basic notions of agency and autonomy, highlighting the fact that they are frequently mistaken by researchers. Briscoe, focuses on the spatial contents of experience and the meaning that "motor system" has for it, he presents Evans' (1982) concept as an alternative to Alva Noë's activism and proposes a concept that is more consistent with current empirical studies (while connecting Evans' concept with Millikan's (1984) research pertaining to mental representations). Petit develops his own interpretation of neurophenomenology (one of the elements of the enactivist programme), focusing on the brain mechanims of consciousness. Ellis also turns towards the functioning of the central nervous system, analysing the role of emotions, self-organisation and anticipation in consciousness, and using this point of view to criticise determinist and epiphenomenalist concepts of consciousness. Gallagher and Bower, while trying to make enactivism more embodied, point towards the important role of emotions and social factors; moreover (as with the two previous authors) they consider the way in which the functioning of the brain should be interpreted in the light of enactivism, especially in the context of research on predictive cognition. Some works consider the applied side of enactivism. The neurophenomenology mentioned above (see article by Petit) is one of the applications. Li and Winchester reflect on the concept of Freedom Education, showcasing both its advantages and its potential cohesion with the enactivist programme. In his interesting, historically grounded essay, Reid considers the relationship between enactivism and theories of teaching, as well as the potential pros and cons of enactivism in these theories, especially in concepts pertaining to teaching mathematics.

In a broader sense, the abovementioned articles can be significantly complemented by two interviews: one with Shaun Gallagher, and the other with Robert Rupert; they are able to direct the attention of the Readers towards more broadly conceived issues connected with situating cognitive processes.

This concludes a broad overview of the current issue. Before we invite you to read the articles collected herein, we will allow ourselves a few more remarks. We can sometimes notice certain continuities in the presented set of texts. One of the continued threads is the criticism of basic notions of enactivism. Aizawa ponders the direction for development and the character of the enactivist revolution. His attention focuses on whether the concept enactivists consider to be cognition should be called cognition at all. He claims that what we have to make do with here is something that was classically considered behaviour, and that makes the status of enactivism in cognitive studies especially interesting. According to the classical approach, it was behavior that was explained (*explanandum*), while the theories of cognition were supposed to explain this behaviour (they were *explanantia*). Enactivists—as Aizawa writes—on the one hand, consider cognition to be *explanandum*, and, on the other, reduce cognition to a form of behaviour. As a result, we have to make do with a particular, cognitive-enactive mixture of *explananda* and *explanantia*, declaratively incompatible, but at the same time, in fact and quite perversely compatible with classical cognitive science. Steiner critically presents the issue of relations between enactivism and representational concepts of cognition. In his critical reflections, Cummins reaches for the basic notions of the first wave of enactivist research. These works, together with Petit's and Bielecka's texts, point towards the weaknesses of enactivist proposals. As we believe, such approaches and their consequences are indispensable in order for enactivism to be able to develop fully and overcome the problems it encounters. Both internal and external criticism are not to be underestimated here[2].

Gallagher and Bower—authors deriving from phenomenological backgrounds—consider the role of the nervous system in discussing enactivism in general. A similar path is pursued by Petit, who discusses one of the research programmes of enactivism: neurophenomenology, as well as by Ellis, who criticises the determinist concepts of consciousness. Such a strong interest in the central nervous system seems to be both surprising and valuable, as it makes possible the meeting of enactivism with the main trend of neurocognitive studies, as well as a reflection (in the light of these studies) on the participatory (Gallagher and Bower) and anticipating (Ellis) concepts of CUN, and also the role of time synchronizations in consciousness (Petit).

In the aforementioned articles we find references to Varela and Maturana's version of enactivism (Aizawa, Cummins, Steiner, Li & Winchester, Ellis), as well as to the enactivist (or activist) proposals of Alva Noë (Briscoe), to Ellis and Newton's framing of enactivism (Ellis), enactivism as proposed by Daniel Hutto (Aizawa) or enactivism in robotics (Bielecka). It is, however, visible that the role of the beginnings of the idea of enactivism remains not to be underestimated, as although the works we have collected frequently refer to very advanced, recent studies, they do not—as we can see—disregard the roots that can undoubtedly be located in Varela and Maturana's research.

Referring back to the quote that opens the present introduction, we can say that thanks to the aforementioned articles we can, in a certain way, get a hold on enactivism and face up to it, and, as a consequence, also to the problem of cognition. We believe that these works can (and should) constitute an important voice in the dispute over what the essence of enactivism (and also enaction) is, and what role it should play in research on cognition. This seems to be suggested both by the critical and the favourable framings of enactivism.

[2] There have already appeared a number of voices criticising enactivism. In our opinion, among the most important of these is the review of Alva Noë's first book written by Ned Block (2005) – a review that, among its other features, charged enactivists (here: Noë) with crypto-behavioralism, as well as mistaking causality for constitution. Additionally, there are interpretation problems concerning basic notions, such as sensomotorical contingencies. In one of her recent texts, Frederique de Vignemont attacks enactivism (which she equates with sensorimotor approaches, as Noë also did), arguing with the claim of the enactivists that every experience constitutes certain particular sensorimotor laws). According to this researcher, it is possible to separate such experiences which—as it seems—cannot be connected with any laws of this kind (see: de Vignemont 2011, 2014).

We hope that the contents of this volume will bring us closer to explaining at least some doubts referenced here and to realising what constitutes cognition according to enactivists—although they themselves believe that it is something as dynamic, momentary, and unstable as handshaking.

References

Aizawa, K. 2014. What is this cognition that is supposed to be embodied? *Philosophical Psychology* (ahead-of-print): 1-21.

Block, N. 2005. Review of Alva Noë, 'Action in Perception'. *Journal of Philosophy*, 102: 259–272.

De Vignemont, F. 2011. A mosquito bite against the enactive approach to bodily experiences. *Journal of Philosophy*, 108 (4): 188-204.

De Vignemont, F. 2014. Acting for bodily awareness. L. Shapiro, ed. *Routldege Handbook of embodied cognition*. Routldege: 287-295.

Evans, G. 1982. *The varieties of reference*. J. McDowell, ed. Oxford: Clarendon Press.

Froese, T., McGann, M., Bigge, W., Spiers, A., & Seth, A. K. 2012. The enactive torch: A new tool for the science of perception. *Haptics, IEEE Transactions on*, *5*(4): 365-375.

Gallagher, S. 2013. The Enactive Hand. *The Hand, an Organ of the Mind: What the Manual Tells the Mental*. MIT Press: 209-225.

Millikan, R. G. 1984. *Language, thought, and other biological categories: New foundations for realism*. MIT Press.


articles

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0002

The Enactivist Revolution

Kenneth Aizawa
Department of Philosophy
Rutgers University, Newark
ken.aizawa[]gmail.com
Received December 2013; accepted June 2014; published Autumn 2014.

Abstract

Among the many ideas that go by the name of "enactivism" there is the idea that by "cognition" we should understand what is more commonly taken to be behavior. For clarity, label such forms of enactivism "enactivism$_b$." This terminology requires some care in evaluating enactivist$_b$ claims. There is a genuine risk of enactivist and non-enactivist cognitive scientists talking past one another. So, for example, when enactivists$_b$ write that "cognition does not require representations" they are not necessarily denying what cognitivists claim when they write that "cognition requires representations." This paper will draw attention to instances of some of these unnecessary confusions.

Keywords: enactivism; enaction; cognition; behavior; autopoiesis.

In soliciting contributions to this special issue of *Avant*, the editors asked whether enactivism fits cognition. This question, however, may well misinterpret what at least some forms of enactivism are about. It may underestimate the breadth of the revolution that at least some strains of enactivism are championing. For some in the enactivist movement, it appears that the goal is not merely to provide a revolutionary new account of what cognition is. We need no more of *that*. What is needed, instead, is a cognitive science that studies something else. What is needed is a cognitive science that does not study cognition! That would be a real revolution.

But, what, one might ask, would such a really revolutionary cognitive science study, if not cognition? One popular proposal is that it should study what has been, and generally continues to be, known as (a type of) behavior.[3] For con-

[3] Another way of making the present point might be to distinguish cognition$_e$ (for enactivist cognition) and cognition$_c$ (for cognitivist cognition). This might make it easier to see that this issue is

venience, and in order to distinguish this form of enactivism from other forms, let us label it "enactivism$_b$," where the subscript indicates the focus on (a type of) behavior. To be sure, not all those who think of themselves as enactivists are enactivists$_b$. Nevertheless, there are prominent representatives of enactivism$_b$. To take one salient example, many enactivists$_b$ have been inspired, at least to some degree, by Humberto Maturana and Francisco Varela's *Autopoiesis and Cognition*. Maturana is relatively explicit about understanding a cognitive system as a (self-maintaining) behaving system. According to Maturana, "A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself, and the process of cognition is the actual (inductive) acting or behaving in this domain." (Maturana and Varela 1980: 13). To take a more recent example, Tony Chemero proposes, "cognitive scientists ought to try to understand cognition as intelligent behavior" (Chemero 2009: 25). Further, he believes that "radical embodied cognitive science can explain cognition as the unfolding of a brain-body-environment system" (Chemero 2009: 43). But, "the unfolding of a brain-body-environment system" sounds like a metaphor for behavior.

The foregoing point might be made in another way. Notice that enactivists$_b$ often propose their theory as an alternative to cognitivism, but they are not always explicit about what in the cognitivist view they reject. As one example, Stewart, Gapenne, and Di Paolo write,

> The aim of this book is to present the paradigm of enaction as a framework for a far-reaching renewal of cognitive science as a whole. There have been many critiques of classical, first-generation cognitivism based on the Computational Theory of Mind. A distinctive feature of this book is a deliberate choice not to go over that old ground yet again, but to reserve the energy for positive exploration of new paths. (Stewart, Gapenne, and Di Paolo 2010: vii).

Di Paolo, Rohde, and De Jaegher 2010, provide a similarly opened-ended rejection of cognitivism:

> Almost two decades since the publication of *The Embodied Mind* (Varela, Thompson, and Rosch 1991), the term enactive has moved out of relative obscurity to become a fashionable banner in many regions of cognitive science. … Theirs was not only an achieved synthesis of existing criticisms to a predominantly computationalist paradigm, but also the articulation of a set of postulates to move these ideas forward. Indeed, the increasing use of enactive terminology serves as an indication that the time is ripe for a new era in cognitive science. To a great extent, we believe this to be so. (Di Paolo, Rohde, and De Jaegher 2010: 33)

---

not about the "right" way to use "cognition" or who gets to use it how. Those who prefer to use this terminology are free to use it, but this paper will adhere to the more mainstream usage. The important point, of course, is not ultimately about the terminology, but the fact that cognitivists and some enactivists are talking about two different things.

One might well think that enactivism$_b$ means to displace the cognitivist's computation theoretic apparatus of rules and (especially) representations with another sort of apparatus. This alternative apparatus might be the mathematics of dynamical systems theory or one or another definition of autopoiesis. In such a vision, enactivism$_b$ and cognitivism are competing theories of the same thing in the way that Newton's theory of gravitation and Einstein's theory of general relativity were competing theories of a single putative force in nature, namely, gravity. This, however, apparently underestimates just how sweeping a change the enactivists$_b$ wish to make in cognitive science. Enactivists$_b$ generally propose to walk away from the issues and concerns of mainstream cognitive science to focus on what is commonly understood as behavior. In practice, therefore, enactivists$_b$ use different tools to study different issues.

While there are times when enactivists$_b$ enthusiastically embrace the dramatic changes implicit in their work, they also tend to paper over the significance of the proposal to study (a type of) behavior by calling behavior "cognition" or "lower-level cognition" or "basic cognition" or "minimal cognition." (See, for example, Calvo and Keijzer 2009, Chemero 2009, Stewart 2010, Di Paolo, Rohde, and De Jaegher 2010, and Hutto and Myin 2013.) This terminology is likely to be misleading to mainstream cognitive scientists, but a more serious problem is that it seems to mislead even some enactivists$_b$. There are times when they write as if they intend to address traditional, sometimes longstanding, problems surrounding cognition. Yet, because they use "cognition" as a term for (a type of) behavior, they are thereby not talking about the same thing as are the traditional cognitivists. Thus, they sometimes fail to come to grips with traditional issues in cognitive science.

This paper will begin, in section 1, with a brief review of the distinction between cognition and behavior as it has formerly been used in cognitive science, namely, that cognitive processes have been thought to be among the many endogenous factors that contribute to the production of behavior. The point here is not to offer definitions of "cognition" or "behavior" or to offer much in the way of clarification of what each of these is, but simply to draw attention to what has been a widely held understanding of the difference between the two. Successive sections (sections 2-5) will then review ways in which Stewart 2010, Froese, Gershenson, and Rosenblueth 2013, and Hutto and Myin 2013, seem not to appreciate the significance of their departure from traditional problems of cognition. Section 6 will emphasize the fact that not all enactivists are enactivists$_b$ by providing clear examples of enactivists who offer a more traditional conception of cognition as a species of endogenous cause of behavior.

As a final preliminary note, it is important to bear in mind that the goal of this paper is clarification, not criticism. Enactivism$_b$ clearly represents a dramatic break with tradition. Enactivists have made that abundantly clear. What bears greater attention, however, is the character of this break. Rather than examine endogenous causes of behavior, such as cognition, enactivism$_b$ proposes to focus on (a type of) behavior. What also bears attention are some of the ramifications of this break. Insofar as enactivism$_b$ no longer addresses cognition as it has formerly been understood, it just so far threatens to ignore cognition. Enactivism$_b$, thus, does not so much solve traditional problems, as merely walks away from them. This, of course, does not bear directly on the truth of enactivism$_b$. It only suggests that enactivists$_b$ need to be more careful in how they deal with traditional problems. If they want to talk about traditional cognition, they apparently need an account of endogenous influences on behavior. Alternatively, if they wish to break with tradition, then they must be careful to make a cleaner break. So, to repeat, the goal of this paper is not so much criticism of enactivism$_b$ as clarification.

1. Cognition and Behavior

One way in which to appreciate the core theoretical commitments of traditional cognitive science might be to revisit some of its founding documents, wherein the original commitments are articulated. As an illustration of this method, Aizawa 2014, describes a bit of common ground between B. F. Skinner and Noam Chomsky, namely, both believed that cognition was a putatively explanatory causal factor in the production of behavior. Where Skinner and Chomsky differed, of course, was in their assessment of the genuine explanatory value of the cognitive. (See, for example, Skinner 1957, and Chomsky 1959.) Another paper that illustrates the core theoretical commitments of traditional cognitive science is the seminal 1958 paper by Alan Newell, J. C. Shaw, and Herbert Simon, "Elements of a Theory of Human Problem Solving." They propose that

> Questions about problem-solving behavior can be answered at various levels and in varying degrees of detail. The theory to be described here explains problem-solving behavior in terms of what we shall call *information processes*. If one considers the organism to consist of effectors, receptors, and a control system for joining these, then this theory is mostly a theory of the control system. It avoids most questions of sensory and motor activities. (Newell, Shaw, and Simon 1958: 151).

This brief passage contains a number of ideas that are relevant to understanding the differences between the traditional information processing approach in cognitive science and enactivism$_b$. These features are worth reviewing in detail.

The very first sentence proposes to treat problem solving as a behavior. Problem solving is, thus, not itself information processing or cognitive processing. In the literature on embodied and enactive cognition, one sometimes encounters the idea that problem solving is not behavior, but is instead cognitive processing. This claim is more complicated than one might expect. It contains a subtle ambiguity. There is one sense of this claim that is entirely uncontroversial and consistent with the Newell, Shaw, and Simon perspective. It seems perfectly reasonable to claim that the entire process of, say, physically manipulating pencil and paper to solve a cryptarithmetic problem counts as problem solving and cognitive processing. This is the sense in which the whole of the process is cognitive processing in virtue of the fact that an important or salient component of the process is cognitive processing. The whole of the manipulative process is cognitive processing, even though strictly speaking only a proper part of the process is cognitive processing. The idea here might be understood through an analogy. The whole of the process of baking a cake, one might say, is not strictly speaking a matter of baking a cake. The process of baking a cake might include breaking some eggs and the mixing of ingredients, processes that are not themselves baking processes strictly speaking. Similarly, the whole of the process of filling up one's car is not strictly speaking a matter of pumping gasoline into the tank. It includes such things as slowing the car, pulling it into the station, and shutting off the engine. In contrast to this unproblematic claim there is the idea that, strictly speaking, the whole of the process of manipulating the pencil and paper is cognitive processing. This would be the sense in which the whole of the process of baking a cake is literally the baking of a cake or the whole of the process of filling one's gas tank is pumping gasoline into the tank. What probably obscures the ambiguity in the claim that problem solving is cognitive processing is the relative lack of clarity about the character of the component processes. There is a relatively clear distinction between slowing the car to pull it into the gas station and pumping the gasoline into the tank, but it is less clear how to distinguish the information processing that might take place only in the brain and what might be called the information processing that takes place in the pencil and paper. It is, therefore, useful in discussing such cases to be clear on the strength of the claim that problem solving is cognitive processing. The claim is subtler than one might have expected.

Second, the passage from Newell, Shaw, and Simon treats behavior as distinct from cognitive or information processing. Information processing is taken to be a mechanism realized in the brain. This flatly contradicts the enactivist$_b$ idea that cognition is (a type of) behavior.

Third, Newell, Shaw, and Simon propose that problem-solving behavior might be explained, in part, by appeal to information processing. On this model, behavior is the thing to be explained, whereas information processing is among the factors that do the explaining. They repeat this idea more emphatically a bit later in their paper: "At this level of theorizing, *an explanation of an observed behavior of the organism is provided by a program of primitive information processes that generates this behavior*" (ibid.) This is a conception of information processing/cognition they shared with Chomsky. Further, it is a conception Skinner recognized as the mainstream conception, arguing, however, that it is misguided.[4] Notice that, by proposing that behavior is distinct from information processing and that information processing is realized in the brain, Newell, Shaw, and Simon implicitly adopt what is sometimes described as the framework of "mechanistic explanation." They propose to explain the behavior of a whole organism primarily by appeal to the behavior of one of its components. This picture might be illustrated with a well-known image from Craver 2007. (See Figure 1.) In this scheme, S ψ-ing would be something like a participant in an experiment solving a problem, whereas, say, x_3 ϕ_3-ing would be the brain processing information.

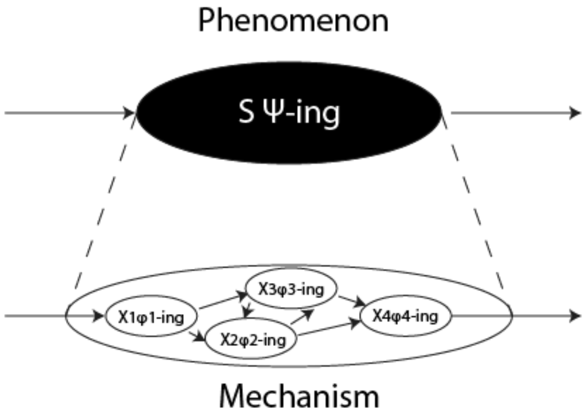


Figure 1. Schema for mechanistic explanation.
Redrawn from Craver 2007: 8, Figure 1.1.

Fourth, and finally, Newell, Shaw, and Simon embrace the traditional cognitive science focus on the role of information processing in the production of behavior, but they do not deny that there can be a role for sensory and motor activities in the production of behavior. We can describe this view by reference to the Craver schema. In the figure above, x_1 ϕ_1-ing might be the eye performing a saccade, where x_4 ϕ_4-ing might be writing with a pencil. Thus, they recognize that there are many component processes that conspire in the production of behavior, but indicate that they will limit their attention to a subset of these factors. The focus of their attention is methodological, not theoretical. In other words, even some of the earliest advocates of information processing psychology anticipated a day when psychologists might

[4] For a contemporary articulation of this picture, there is the *Stanford Encyclopedia of Philosophy* entry on cognitive science. (Thagard 2010).

take up the issues concerning the contributions of sensory and motor factors to the production of behavior.

This traditional picture of the relationship between cognition and behavior—that cognition is among the factors that might explain behavior—is rarely, if ever, explicitly discussed in the enactivist literature. Nevertheless, if one maintains that cognition is (a type of) behavior, one seems to be walking away from much of what cognitive science has been up to, namely, the study of putative endogenous contributions to the production of behavior. Nevertheless, enactivists$_b$ have often seemed willing to do this, suggesting that a fresh start for cognitive science is in order. Setting aside questions about the wisdom of walking away from so much of the work that has been done in cognitive science, this paper will show some of the missteps this has engendered.

2. Bourgine and Stewart 2004

Recall Maturana's claim that "A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself, and the process of cognition is the actual (inductive) acting or behaving in this domain." (Maturana and Varela 1980: 13). It is not entirely clear what is going on in this brief passage. Maturana apparently claims that the process of cognition is behaving in a domain—that cognition is a type of behavior—but one might have one's doubts that he genuinely proposes to identify cognition and behavior. How is this conception of cognition supposed to relate to the traditional conception, if at all? And, what are we to make of the character of this claim? Is it supposed to be a definition, a conceptual analysis, a theoretical hypothesis, or something else? It might well be read as a stipulative definition, but then again there are times when Maturana is prone to forceful pronouncements about empirical matters. Matters here are not that clear.

Some of the ambiguity in Maturana's text is eliminated from the account in Bourgine and Stewart 2004. The latter presentation is more deliberate and explicit about *defining* cognition as behavior and that this definition does not capture what is "ordinarily" meant by "cognition." Their proposal, therefore, at least looks more like a stipulative definition of "cognition." To provide their definition, they, first, define A interactions as those system-environment interactions that have consequences for the internal state of an organism and B interactions as those system-environment interactions that have consequences for an organism's immediate environment or modify the relation of the organism to its environment. These terms then figure into a definition of "cognition":

> D-C1: A system is cognitive if and only if type A interactions serve to trigger type B interactions in a specific way, so as to satisfy a viability constraint.

Bourgine and Stewart explicitly decline to define what a "viability constraint" is, but the rough idea is that A interactions must trigger B interactions that are "good for" the system. This proposal seems to imply that systems are cognitive when stimuli provoke them to behave in ways that they are "good for" the system. And Bourgine and Stewart subsequently substantiate this interpretation:

> It may be useful to illustrate this by examples of interactions such as falling down stairs, eating, or breathing (including the breathing of a poisonous but odorless gas). Ordinarily, such interactions are not considered as "cognitive." On the definition proposed here, they will not be cognitive unless the consequences for the internal state of the system are employed to trigger specific actions that promote the viability of the system. Thus, falling down stairs will be cognitive if but only if the fall triggers reactions such as a modification of muscle tone that limits the damage; and this does require specific sensory and motor organs. Similarly, eating is cognitive if but only if it triggers a reaction of satiety that prevents damage from overeating; breathing a poisonous gas is cognitive if but only if it triggers evasive action, which will require a specific sensory organ to detect the poison, and the resulting sensation to trigger an appropriate, coordinated motor response. (Bourgine and Stewart 2010: 338).

Bourgine and Stewart evidently concur with the cognitivist view that falling down stairs, eating, and breathing are not ordinarily considered to be cognitive (processes). Yet, they differ from cognitivists in their rationale for this assessment. For Bourgine and Stewart, it is only *some* instances of falling down stairs, eating, and breathing that are not to count as cognitive, namely, those in which there are no prophylactic effects, such as changes in muscle tone or satiety. For cognitivists, however, falling down, eating, and breathing are not, strictly speaking, cognitive processes at all;[5] they are, at most, behavior. Moreover, they are likely to receive distinct behavioral analyses. By cognitivist lights, many instances of eating may count as cognitive behaviors insofar as they require cognitive processes in order to do things such as recognize food or to plan how to use knife and fork to obtain bite-sized pieces of food. By contrast, instances of breathing may not count as cognitive behaviors insofar as they do not involve cognitive processing. Humans typically breathe without thinking about it. Non-cognitive autonomic processes generally suffice for breathing. Finally, relatively few cases of falling down stairs will count as cognitive behaviors. Usually, gravity can do most of the work of tumbling someone down stairs without their really thinking about it.

[5] For clarification of the qualifier, "strictly speaking," recall the first point in the discussion of the passage from Newell, Shaw, and Simon 1958, in section 1 above.

The point here, of course, is not to pin down precise frequencies or analyses of these cases, but merely to give a nod to the sorts of factors that will enter into cognitivist analyses and to contrast them with Bourgine and Stewart's analysis.

As we have just seen, the traditionalist might well claim that Bourgine and Stewart's theory is subject to counterexamples in which they call things that are not cognitive, cognitive. Falling down stairs is not a cognitive process, even though Bourgine and Stewart's theory says it is. A traditionalist can also easily imagine cases in which Bourgine and Stewart's theory would call cognitive processes "non-cognitive." So, imagine a person, Jane, looking out the window and seeing a cloudy day and thinking, "It looks like rain." This looks to be what Bourgine and Stewart would call an A interaction. Jane might then think, "Maybe I should take my umbrella. But, then again, I have a lot to carry today. Maybe I should just chance it and leave my umbrella at home," before she finally walks out the door.[6] Jane's interior monologue and the walking out the door looks to be what Bourgine and Stewart would call a B interaction: it is a system-environment interaction that modifies the relation of Jane to her environment. But, now, suppose that a dramatic cloudburst drenches Jane and that this is not good for her. By Bourgine and Stewart's enactivism$_b$, Jane's interior monologue was not a cognitive process, nor was there anything like a thought process underlying her interior monologue. This looks, to a cognitivist, like a counterexample to Bourgine and Stewart's theory in which a cognitive process is mistakenly labeled "non-cognitive."[7]

At this point, one might propose that Bourgine and Stewart can simply stand by their stipulative definition of what a cognitive process is. Their theory does not capture traditional "intuitions," "theories," or—one might say, "false leads," but that is not what it is meant to do. Bourgine and Stewart, therefore, have a kind of theoretical "safety" in offering a stipulative definition of what

[6] Benny Shannon is an enactivist (though not an enactivist$_b$) for whom the study of such thought sequences is crucially important. (Cf., Shannon 2010.) As will emerge, Bourgine and Stewart's theory of cognition appears to have the consequence that some of the thought sequences that Shannon has studied will not be cognitive. Or, if one prefers, Bourgine and Stewart's theory of cognition$_b$ appears to have the consequence that some of the thought sequences that Shannon has studied will not be cognitive$_b$.

[7] This sort of scenario can be used to draw attention to another feature of Bourgine and Stewart's account. We cannot tell just from the occurrence of the A interaction and the B interaction whether or not a process is cognitive in Bourgine and Stewart's sense. Whether Jane's interior monologue was (indicative of) a cognitive process or not apparently depends on whether or not it actually rains and what impact this has on her. So, if it does not rain and this turns out to be good for Jane, say by sparing her the burden of carrying an unnecessary umbrella, then we have a cognitive process. Alternatively, if a downpour drenches Jane and this is not good for her, then the interior monologue would not be (indicative of) a cognitive process. Bourgine and Stewart's theory makes a process cognitive or non-cognitive based on contingent events that take place after the putative thought process.

they mean by "cognition." The problem with this, however, is that when Stewart tries to use the theory to address traditional problems, he thereby misses his target.

Stewart claims that "There are two basic requirements for any paradigm in cognitive science: it must provide a genuine resolution of the mind-body problem, and it must provide for a genuine core articulation between a multiplicity of disciplines" (Stewart, 2010: 1). He then proposes a solution to the mind-body problem:

> As discussed in Bourgine and Stewart 2004, we may define a system as "cognitive" if and only if it generates its actions, and the feedback sensations serve to guide actions, in a very specific way so as to maintain its autopoiesis and hence its very existence. With these definitions, "cognition" and "life" are fundamentally the same phenomena; and, in principle, the mind-matter problem is solved. (Stewart 2010: 1-3).

Thus, Stewart has proposed to solve the mind-body problem by a form of type identity theory: cognitive processes are biological processes (life processes).[8] Then biological/life processes are, in turn, identified with physico-chemical processes (Cf., Stewart 2010: 203). So, it looks like Stewart and his enactivism$_b$ strike squarely at one of the central issues in the philosophy of mind.

Appearances here are deceiving. If Stewart maintains that by "cognition" he does not mean what has traditionally been meant by cognition—if he does not correctly characterize what has traditionally been meant by "cognition" or "the mind," then he is not addressing the traditional mind-body problem. The traditional mind-body problem has not been concerned with how to relate falling down, eating, or breathing to biological or physico-chemical processes. Instead, the traditional mind-body problem concerns what is perhaps a cluster of problems, none of which centers on behaviors.

Notice that the mind-body problem as found in Descartes' philosophy is a question of how there could be causal interactions between an immaterial soul or mind and a material body. But, this is a question of how cognition, as traditionally construed, can interact with bodily processes. It is not a question of how cognition, construed as behavior, can interact with bodily processes. Stewart's "cognition" does not speak to the Cartesian version of the mind-body problem.

Kim 2005, however, observes that the Cartesian mind-body problem is not the contemporary mind-body problem found in the philosophy of mind and cog-

[8] One might well make the case that Stewart does not have a type identity theory solution but a functionalist solution. The difference probably does not make a difference here, since the weakness in Stewart's purported solution lies in his view that cognition is viable behavior. Maybe there are other enactivist$_b$ tools for dealing with this portion of the traditional mind-body problem, but the Bourgine-Stewart theory of cognition alone will not suffice.

nitive science. Instead, the contemporary mind-body problem appears to be a cluster of problems. One of these is a problem of mental causation. This problem, however, is not (at least in the first instance) about how bodily movements can be causally efficacious; it is about how internal states and processes, such as beliefs, desires, intentions, and decisions, could be among the causes of behavior. (Cf, Kim 2005: 7f). If Stewart is not thinking of cognitive states as internal causes of behavior, then he is not grappling with the mental causation portion of the mind-body problem. A second problem in the cluster of mind-body problems is the problem of consciousness. By consciousness, Kim has in mind qualitative features such as "the smell of the sea in a cool morning breeze, the lambent play of sunlight on brilliant autumn foliage, the fragrance of a field of lavender in bloom, and the vibrant, layered soundscape projected by a string quartet" (Kim 2005: 11). David Chalmers seems to have a similar thing in mind, though described with different examples: "the quality of deep blue, the sensation of middle C" (Chalmers 1995: 201). Insofar as Stewart's theory of "cognition" (or mind) does not capture what is traditionally meant, but constitutes a stipulative definition not tied to these cases, Stewart evidently fails to address the traditional mind-body problem. We might concede, if only for the sake of being agreeable, that Bourgine and Stewart do solve *a version of* the mind-body problem, namely, the version that shows how cognition defined in D-C1 can be related to the body. But, even with that concession, it remains true that Stewart has not resolved the traditional mind-body problem(s). Instead, enactivism$_b$ threatens to leave aside old issues for new issues.

So, to summarize this section, one might begin with the possibility that Bourgine and Stewart have proposed a definition of "cognition" that is meant to capture what has traditionally been thought of as cognition. Two sorts of counterexamples, however, challenge this view. 1) Falling down stairs is not a cognitive process; it is (at best) a cognitive behavior. 2) Interior monologues constitute, or are indicative of, cognitive processes even in cases where these interior monologues (or the processes underlying them) lead to bad outcomes for their possessors. In light of these considerations, it is perhaps better to interpret Bourgine and Stewart as not merely overthrowing the theoretical apparatus of cognitivism with its computational rules and representations. Instead, they also propose to overthrow the putative exemplars of cognition. Cognition is not one among many possible endogenous causal contributors to behavior; it is, instead, viable behavior. If, however, this is what Stewart is up to, then he appears to be walking away from many of the problems that cognitivism faced. For example, Stewart's identification of cognition and life does not solve the traditional mind-body problem. It solves, at best, a novel version of the mind-body problem.

3. Froese, Gershenson, and Rosenblueth 2013

The hypothesis of extended cognition, in at least some of its earliest articulations, maintains that, while brains realize some cognitive processes, sometimes, under certain conditions, larger configurations of brain, body, and world also realize cognitive processes. This is the conception that was implicit in parts of Andy Clark and David Chalmers' seminal paper. (See Clark and Chalmers 1998.) It was the conception in play in the two "cognitive equivalence" arguments based on the play of the video game Tetris and the Inga-Otto thought experiment. Recall the description of three modes of Tetris play:

> (1) A person sits in front of a computer screen which displays images of various two-dimensional geometric shapes and is asked to answer questions concerning the potential fit of such shapes into depicted "sockets". To assess fit, the person must mentally rotate the shapes to align them with the sockets.
>
> (2) A person sits in front of a similar computer screen, but this time can choose either to physically rotate the image on the screen, by pressing a rotate button, or to mentally rotate the image as before. We can also suppose, not unrealistically, that some speed advantage accrues to the physical rotation operation.
>
> (3) Sometime in the cyberpunk future, a person sits in front of a similar computer screen. This agent, however, has the benefit of a neural implant which can perform the rotation operation as fast as the computer in the previous example. The agent must still choose which internal resource to use (the implant or the good old fashioned mental rotation), as each resource makes different demands on attention and other concurrent brain activity. (Clark and Chalmers 1998: 7).

Case (1) looks to be a case of someone playing the video game using old-fashioned, brain-internal information processing. By contrast, (2) and (3) are supposed to be cognitively the same as (1) with the only difference between them being the material substrates that realize cognition. Next recall the Inga-Otto thought experiment. Inga has normal human memory, reads that there is an interesting exhibit at the Museum of Modern Art, and decides that she would like to see it. She thinks for a moment, recalls that it is on 53rd St., then heads on her way. By contrast, Otto is suffering from the early stages of Alzheimer's Disease, so he has developed a system for maintaining information in a notebook. In this notebook, he has the address of the Museum of Modern Art. When he reads of the exhibit at the museum, he decides that he would like to see it. He then picks up his notebook, flips through it until he finds the address, then heads on his way. Clark and Chalmers claim that "in relevant respects the cases are entirely analogous: the notebook plays for Otto the same role that memory plays for Inga. The information in the notebook functions just like the information constituting an ordinary non-occurrent belief; it just happens that this information lies beyond the skin" (Clark and Chalmers 1998: 13).

Whereas Clark and Chalmers sometimes offer "cognitive equivalence arguments" for extended cognition, Froese, Gershenson, and Rosenblueth 2013, propose a different path. They reject the idea that there is ever any brain-bound, information processing type cognition of the sort that Inga was said to have. Instead, they adopt the enactivist$_b$ view that cognition is (a type of) behavior: "cognition is primarily conceived of as a form of viable conduct by an agent in an environment" (Froese, Gershenson, and Rosenblueth 2013: 1). Thus, they take it that "[the Dynamical Systems Hypothesis] takes the notion of an extended mind as its starting point, rather than as a curious exception" (ibid.) Froese, Gershenson, and Rosenblueth, therefore, defend extended "cognition" in the sense of extended adaptive behavior. This path, however, does nothing to help the version of extended cognition in the Tetris and Inga-Otto examples. Through those examples, Clark and Chalmers maintained the bold conclusion that what was once thought to have been realized only in the brain—a type of information processing—is, in fact, also sometimes realized in the brain, body, and world. Thus, Froese, Gershenson, and Rosenblueth, do not so much defend the hypothesis of extended cognition as leave it in favor of the hypothesis of extended viable conduct.[9]

Suppose, then, for the sake of argument that we follow Froese, Gershenson, and Rosenblueth and think about viable conduct. If so, then the further step to the view that such "cognition" is extended is trivial. "Conduct" is essentially another word for behavior and behavior typically is realized by brain, body, and world. In cases in which hammering in a nail is a viable behavior, it is apparently realized by cognitive, attentional, and motivational processes in the brain, along with the propagation of light in the eye, along with muscular processes in the arms, and contact between the hammer and a nail. Who would doubt that? Why would anyone doubt that? So, where the hypothesis of extended cognition is a controversial hypothesis, the hypothesis of extended viable conduct appears to be widely if not universally accepted. What makes the latter *appear* interesting is the enactivist$_b$ terminology of using "cognition" as a term for viable conduct, where most cognitive scientists think of "cognition" as a term for cognition.

Froese, Gershenson, and Rosenblueth appear to appreciate that they cannot rely on the view that cognition is a form of viable conduct in order to argue for the extended cognition hypothesis. Therefore, they argue that "even if we accept [the] internalist starting point, a proper understanding of neuronal activity will force us to accept an extended view of the mind nonetheless." (Froese, Gershenson, and Rosenblueth 2013: 2). Given this concession, they propose two distinct arguments for extended cognition. There is an extensive

[9] If one prefers, one might say that Clark and Chalmers defend something more like extended cognition$_c$, where Froese, Gershenson, and Rosenblueth defend extended cognition$_b$. Again, the body of the text uses more standard terminology.

discussion of an evolutionary robotics simulation, but the details of that are irrelevant to their arguments, so will not be reviewed.

Consider their two arguments in reverse order of their appearance. Their second argument is relatively simple:

> 6. The non-isolated [Continuous-Time Recurrent Neural Network]'s output is determined by its input, albeit mediated by its internal activity, while this input is determined by its motor output, albeit mediated by bodily and environmental (including social) activity.
>
> 7. It logically follows from the above that the non-isolated CTRNN's additional neural complexity is partially constituted by its own sensorimotor and social coupling. (Froese, Gershenson, and Rosenblueth 2013: 6).[10]

This is a relatively simple version of the infamous "coupling-constitution fallacy".[11] Froese, Gershenson, and Rosenblueth's contention notwithstanding, the second sentence above does not logically follow from the first. In the most simplistic of coupling arguments, one might observe that a cognitive process X is causally influenced by a *prima facie* non-cognitive process Y, then infer that the appearances of these processes are misleading. Instead, the entire Y-X process is a cognitive process. The Froese, Gershenson, and Rosenblueth example, however, is a bit more complicated, since there are supposed to be reciprocal causal connections wherein Y causally influences X and X causally influences Y.

Reciprocal causal connections, however, do not suffice to close the gap between causation and constitution. Suppose we accept the internalist view that the non-isolated CTRNN has some "neural complexity." The premise in 6 notes that the CTRNN output is causally determined by its input and its input is causally determined by its output. This is a causal premise. The conclusion, however, is that the "additional neural complexity" is constituted by the processes outside of the CTRNN. But, why go beyond thinking that the additional neural complexity is merely causally influenced by processes outside the CTRNN? There seems to be no warrant for the conclusion that the additional neural complexity is constituted by external processes.[12] Note that, in a nor-

[10] Technically, the argument seems to be for extended "neural complexity," whatever that is. To make this relevant to the hypothesis of extended cognition, there would need to be some link between "neural complexity" and cognition and it is unclear what Froese, Gershenson, and Rosenblueth propose this link to be. They think that cognition is viable conduct, not neural complexity. And, traditional cognitivism takes cognition to be something like rule-governed, symbol manipulation or information processing, not "neural complexity." For present purposes, however, these idiosyncrasies can be set aside.

[11] Cf., e.g., Block 2005, Adams & Aizawa 2008, Rupert 2009.

[12] Froese, Gershenson, and Rosenblueth seem to think that it matters whether the environment qualitatively changes the CTRNN implementing an artificial nervous system (ANS):

mally functional HVAC system, the thermostat's output is determined by its input, albeit mediated by its internal activity, while its input is determined by its output, albeit mediated by environmental activity, such as the burning of natural gas in the furnace. Nevertheless, the change of shape of the bimetallic strip in the thermostat is still limited to the bimetallic strip.

Froese, Gershenson, and Rosenblueth's second argument relies on a distinction between the properties of parts and the properties of wholes and the idea that these properties are often qualitatively distinct. This is an idea that is quite familiar from the literature on mechanistic explanation and is illustrated in Figure 1, in section 1, above. The idea is that entities do things in virtue of their parts doing qualitatively distinct things. The idea also appears in the enactivism literature in the introduction to Hutchins 2010:

> Distributed cognition is a framework for exploring the cognitive implications of the commonsense observation that in systems characterized by multiple levels of interacting elements, different properties may emerge at different levels of organization. Thus, a colony of social insects has different properties than any individual insect in the colony. At the level of organisms, bodies have different properties than organs, which have different properties than cells. In the realm of cognition, a neural circuit has different properties than any of the neurons in the circuit. The same can be said of a brain area with respect to the neural circuits that compose it, or of an entire brain with respect to the areas that interact within the brain. This is also true of the body/brain system with respect to either brain or body, and the world/body/brain system with respect to any of its parts. A system composed of a person in interaction with a cognitive artifact has different cognitive properties than those of the person alone. (Hutchins 2010: 425.)

Froese, Gershenson, and Rosenblueth, propose to avoid the coupling-constitution fallacy by appeal to something like this picture:

> This critique is known as the "coupling-constitution fallacy". In this paper we respond to this reductionist challenge by using an evolutionary robotics ap-

---

> our aim is to build a model of an embodied agent, whose artificial nervous system (ANS) has mathematical properties that are in principle impossible for it to have in isolation. The motivation for this criterion is the need to go beyond a demonstration of how an agent's situatedness within a sensorimotor loop modulates the internal activity of the ANS, but can transform the ANS into a qualitatively different kind of system altogether.
>
> if an ANS with less than 3D is nonlinearly coupled with other non-chaotic systems, and its internal neural activity spontaneously becomes chaotic, then an explanation of this property as resulting from an extended process of interaction cannot be accused of committing the coupling-constitution fallacy. (Froese, Gershenson, and Rosenblueth 2013: 2).

So, their idea is that when this ANS interacts with the environment, it undergoes a qualitative shift. It becomes chaotic. But, what does this have to do with cognition? It isn't that becoming chaotic is the same as becoming cognitive, is it? And, what does this have to do with what does, or does not, constitute a cognitive process? Why not stick with the idea that interaction with the environment causes the ANS to become chaotic?

> proach to create a minimal model of two acoustically coupled agents. We demonstrate how the interaction process as a whole has properties that cannot be reduced to the contributions of the isolated agents. We also show that the neural dynamics of the coupled agents has formal properties that are inherently impossible for those neural networks in isolation. (Froese, Gershenson, and Rosenblueth 2013: 1).

> From the perspective of the DSH, which proposes a distributed view of cognition as the default mode of cognition, there is no coupling-constitution fallacy because properties of the sensorimotor interaction process cannot be reduced to that of the isolated components. (Froese, Gershenson, and Rosenblueth 2013: 6).

> It is a good strategy for the EMH to highlight that ongoing interaction between a cognitive agent and environment results in a novel, mutually encompassing process with new properties of its own. (Ibid.)

With this picture, the work-around for the coupling-constitution fallacy is easy. There are brain properties, bodily properties, environmental properties, and brain-body-environment properties. In other words, there are properties the brain has, properties the body has, properties the environment has, and properties a brain-body-environment system has. One does not need to say that it is interaction between brain, body, and world that converts bodily properties and environmental properties into new properties. Instead, a causal interaction between brainy, bodily, and environmental processes gives rise to a new, qualitatively distinct, emergent property that is cognitive. Problem averted.

Notice that, for this argument, one does not really need a whole lot of evolutionary robotics simulation. It depends primarily on the picture of mechanistic explanation according to which higher level properties are realized by properties of lower level individuals. All of that is fine. There is, however, one tacit premise that is the downfall of this reply to the coupling-constitution fallacy. Recall that Froese, Gershenson, and Rosenblueth proposed to begin with the internalist assumption that cognitive processes are realized by the brain. Recall that they wanted to show that "even if we accept [the] internalist starting point, a proper understanding of neuronal activity will force us to accept an extended view of the mind nonetheless." (Froese, Gershenson, and Rosenblueth 2013: 2). But, if they concede that assumption, namely, that the brain properties are cognitive, then given their other premise that the properties of the brain-body-world system are qualitatively distinct from the properties of the brain, the body, and the world, then this gives us the beginnings of an argument that the properties of the brain-body-world system are not cognitive! The argument potentially backfires.

Of course, Froese, Gershenson, and Rosenblueth may be too quick to concede that brain properties are cognitive. Perhaps they should simply stick to their view that cognitive processes are instances of viable conduct. If they do that, however, then there really is no need for the coupling kinds of arguments at all. Clearly viable conduct is extended. As noted above, no one doubts that when hammering in a nail is a viable behavior, it is probably realized by cognitive, attentional, and motivational processes in the brain, along with the propagation of light in the eye, along with muscular processes in the arms, and physical processes in the hammer, nail, and wood. Once you have the view that cognition is a type of behavior, it is relatively smooth sailing to the conclusion that such "cognition" is extended.

So, the upshot of our discussion of Froese, Gershenson, and Rosenblueth's paper is three-fold. If they wish to defend the view that viable conduct is extended, then they are doing nothing to support one of the original versions of the idea that cognition is extended. They are not at all supporting the idea that some sort of information processing style of cognition is extended. Second, if they wish to defend the view that viable conduct is extended, then they are not defending a view that it seems anyone has ever doubted. The standard view in cognitive science is that conduct or behavior is extended. Third, the appeal to the framework of mechanistic explanation does nothing to avoid the problems of the coupling-constitution fallacy, unless one begins with the assumption that cognition is a property of a brain-body-world system. But, if one has that assumption, there is no need for further argumentation using the framework of mechanistic explanation. Properties of a brain-body-world system are clearly extended. What this suggests is that, by taking "cognition" to be a term for viable conduct, rather than for some endogenous causal contributor to the production of conduct, Froese, Gershenson, and Rosenblueth have marginalized their view from the concerns of extended cognition and the coupling-constitution fallacy.

4. Hutto and Myin 2013

From the earliest pages of their *Radicalizing Enactivism*, Daniel Hutto and Erik Myin challenge the view that all cognition requires representations. Their alternative is Radically Enactive (or Embodied) Cognition (REC):

> The most radical versions of these approaches are marked by their uncompromising and thoroughgoing rejection of intellectualism about the basic nature of mind, abandoning the idea that all mentality involves or implies content. Call this—the view we defend—Radically Enactive (or Embodied) Cognition—REC for short. (Hutto and Myin 2013: 1)

> Embodied ways of thinking reject the familiar explanatory framework of orthodox cognitive science in favor of alternative platforms. Adherents of such views deny that the best way to explain cognition is to posit the construction of internal representational models (ibid.: 2)

> Defenders of REC argue that the usual suspects—representation and computation—are not definitive of, and do not form the basis of, all mentality. (ibid.: 3).

Notice in these claims they write of rejecting "the familiar explanatory framework of cognitive science" and "the usual suspects—representation and computation." The familiar view, however, is that all cognition, understood as one of the internal, brain-realized causes of behavior, involves representation. The tradition does not, however, maintain that (adaptive) behavior involves representation. Plants, for example, might display adaptive behaviors, such as phototropism, without deploying representations to do this. Such cases would be cases in which plants produce behaviors that are not cognitive behaviors. They are not behaviors that are produced, in part, through cognitive processes.

By contrast, when Hutto and Myin stake out their own view about cognition, they apparently have a different conception of cognition. They use "basic cognition" as a phrase to describe what sounds like behavior: "We restrict our ambitions to promoting REC, calling upon strong versions of two theses. We dub these the Embodiment Thesis and the Developmental-Explanatory Thesis. The former equates basic cognition with concrete spatio-temporally extended patterns of dynamic interaction between organisms and their environments." (Hutton and Myin 2013: 5). For Hutto and Myin, "basic cognition" is concrete spatio-temporally extended patterns of dynamic interaction between organisms and their environments. This is a description of what traditional cognitive science would call "behavior." They also seem to use "cognition" as a term for behavior. In the early pages of their book, they write that "proponents of Enactive and Embodied ways of thinking reject the familiar explanatory framework of orthodox cognitive science in favor of alternative platforms. Adherents of such views deny that the best way to explain cognition is to posit the construction of internal representational models built on the basis of retrieved informational content." (Hutto and Myin: 2). It is somewhat odd to say that traditional approaches explain *cognition* by positing the construction of internal representational models. As we saw with Newell, Shaw, and Simon, the traditional approach is to explain *behavior* by positing representations, among other things. This oddity disappears, however, if we understand them to use "cognition" as a word for behavior. Second, they claim that "Enactivism is inspired by the insight that the embedded and embodied activity of living beings provides the right model for understanding minds." (ibid.: 4). Embedded and embodied activity, which sounds like another description of behavior, probably would be a very good model for the mind, if the mental were the behavioral.

The stage is now set for traditional cognitive science to go about its business of studying cognition that purportedly must involve representations, where enactivism$_b$ goes about its business of studying behavior which need not involve representations. Nevertheless, things do not go so smoothly. Hutto and Myin's use of the phrase "basic cognition" muddies what might otherwise be a very simple argument. Hutto and Myin want to challenge the standard view that all cognition requires representation and one way they wish to do this is by providing counterexamples. As a point of logic, they evidently want a case, or two, in which there is cognition without representation. But, instead of cases in which there is cognition without representation, they only provide cases of their "basic cognition" without representation. That is, they only provide cases of behavior without representation.

Hutto and Myin's two putative counterexamples are Rodney Brooks' behavior-based robots and Barbara Webb's models of cricket phonotaxis. (See Brooks 1997, Webb 1994, 1996.) Even as Hutto and Myin describe the example, Brooks' robots show only cases of behavior that does not require representation: "Brooks' first-generation behavior-based robots, and those that followed, succeed precisely because their behaviors are directly guided by continuous, temporally extended interactions with aspects of their environments, rather than being based on represented internal knowledge about those domains (knowledge that would presumably be stored somewhere in the robots' innards)." (ibid.: 42). Such an analysis is irrelevant to mainstream cognitive science, since it only shows that behavior does not require representation, not that cognition does not require representation. Thinking of "cognition" and "behavior" as referring to the same thing seems to mislead Hutto and Myin. The same oversight appears in their analysis of Webb's models of crickets. By their own analysis, "Webb's (1994, 1996) work on cricket phonotaxis offers a vivid example of a model of how successful navigation takes place in the wild, apparently without the need for representations or their manipulation. … In other words, the capacity of these animals to adjust their behavior when successfully locating mates requires them to engage in a continuous interactive process of engagement with the environment." (ibid.: 43). So, by Hutto and Myin's own analysis, these are only cases of behavior without representation; not cases of cognition without representation.[13]

What the foregoing suggests is that enactivist$_b$ terminology confuses even Hutto and Myin. It complicates what should be a simple exercise is trying to generate a counterexample to the claim that all cognition requires representation. Rather than offering examples of cognition that do not involve representation, they only offer examples of behavior that do not involve representation. This, of course, only shows that one argument for REC has not worked. It does not show that the argument cannot be fixed, as by revisiting the examples and

[13] AUTHOR shows how Chemero 2009, makes much the same sort of mistake in thinking that models in Beer 2003, and van Rooij, Bongers, and Haselager 2002, "show how radical embodied cognitive science can explain cognition as the unfolding of a brain-body-environment system, and not as mental gymnastics" (Chemero 2009: 43).

showing how they have internal mechanisms that are plausibly cognitive, but which nevertheless contain no representations. Nor does it show that there are no other arguments for REC that might work.[14] Much less does it show that REC is false. As billed at the outset, the goal here is to clarify some of missteps in the enactivist$_b$ revolution. The point of clarification here is that Hutto and Myin's formulation of an argument based on the Brooks and Webb models do not work as billed.

5. Not all Enactivists are Enactivists$_b$

The goal of this paper has not been to critique the whole of the enactivist program. As is often noted, enactivism is still in its formative stages and many ideas remain to be worked out on many different topics. Some authors associated with enactivism do not seem to be concerned with cognition at all. So, for example, Barbaras 2010, which was included in Stewart, Gapenne, and Di Paolo's *Enaction* anthology, wrote about life and metabolism and barely mentioned cognition. Moreover, as is to be emphasized now, not all enactivists are enactivists$_b$. One can say this, but its force might be better appreciated if we describe enactivists who are not enactivists$_b$.

Perhaps one of the more significant examples is that, some years ago, Stewart seems not to have been an enactivist$_b$.[15] In the abstract to a 1996 paper, Stewart writes, "In contemporary cognitive science, there are two distinct paradigms with contrasting definitions of cognition. The computational theory of mind is based on the syntaxical manipulation of symbolic representations; this paradigm is objectivist because the postulate of a unique independent reality is necessary as a referential basis for semantic grounding of the symbols. The alternative 'constructivist' paradigm is based on the biological metaphor 'cognition = life' and programmatically follows the evolution of cognition from bacteria to civilized humans; it is non-objectivist." (Stewart 1996: 311). Rather than articulating two distinct accounts of what cognition is, however, Stewart appears to be thinking of something along the lines of what philosophers might interpret as a distinction between a form of realism and a form of anti-realism: "In an objectivist [computationalist] perspective, cognition is the subjective representation of an ontologically independent objective reality. In a constructivist perspective, based on the biological metaphor 'cognition = life', the clear point of contrast is that the subject and the object of knowledge

[14] In fact, Hutto and Myin spend a lot of time arguing that causal and informational approaches to naturalizing content have failed, hence that we therefore have some reason to think we should abandon the hypothesis that all cognition requires representation. This argument is untouched by the foregoing.

[15] Shannon 2010, also takes an enactivist, but not an enactivist$_b$ position. Limitations of space preclude a discussion of this.

are not independent but are mutually constitutive" (Stewart 1996: 316.)[16] Stewart's "objectivism" and "constructivism" seem to be versions of realism and anti-realism.

In fact, it appears that at two points in this earlier paper Stewart understands cognition in a more traditional way, not as behavior, but as an internal mechanism that contributes to the production of behavior. As the first point, there is his commentary regarding what appear to be (cognitive) mechanisms by which animals can cancel out the effects of bodily movements in order to perceive organism independent objects.

> Animals with a central nervous system have the capacity to distinguish within their own cognitive repertoire between modifications of their sensory input which are the immediate consequence of their own actions, and modifications which are not so caused. For example, when an animal moves its eyes, the retinal image (and hence the stimulation of the retinal cones) is modified, but a mammal does not usually confuse this movement with the movement of an object in its environment. The construction of perceptual invariants on the basis of motor-sensory correlations of this sort is thus at the basis of the emergence of a 'stable external world' populated by 'objects' which exist as such in the cognitive repertoire of the organism itself. … Bacteria (or trees), for example, are quite incapable of cognitive feats of this sort. (Stewart 1996: 320).

Here it is at least possible that Stewart conceives of there being cognitive mechanisms that enable (some) animals, but not bacteria and trees, to avoid confusing the effects of self-movement with the effects of object movement. In fact, these different mechanisms might be just the kinds of differences between (some) animals, on the one hand, and trees and bacteria, on the other, that cognitivists would contend mark the difference between cognitive and non-cognitive processes. As a second point, Stewart notes that

[16] It is indicative of the cross-currents in enactivism that Hutchins 2010, was part of the Stewart, Gapenne, and Di Paolo *Enaction* anthology, but also apparently endorses the kind of computational theory of mind that Stewart would reject as "objectivist". Hutchins describes computational transformations on representations of what appear to be navigator-independent features of the world:

> Two successive positions of a ship are plotted on a three-minute interval. Suppose the distance between them is 1500 yards. The navigator computes ship's speed to be 15 knots by doing the following: "The distance between the fix positions on the chart is spanned with the dividers and transferred to the yard scale. There, with one tip of the divider on 0, the other falls on the scale at a tick mark labeled 1500. The representation in which the answer is obvious is simply one in which the navigator looks at the yard-scale label and ignores the two trailing zeros" (Hutchins 1995…, 151-152). In this analysis, high-level cognitive functions were seen to be realized in the transformation and propagation of representational states. The span between the fix positions on the chart is a representational state that is transformed into a span on the dividers. This representational state is then transformed into a span on the yard scale. Finally, the span on the yard scale is transformed into the answer by reading the label on the designated tick mark in a particular way. (Hutchins 2010: 427).

Hutchins, thus, also appears to be an enactivist, but not an enactivist$_b$.

> Armed with representations of this sort, an organism can set itself a 'goal' (expressed in terms of a desired perceptual configuration), and then by purely mental activity (without having to take the risks involved in proceeding by trial and error by actually acting in the world) elaborate a sequence of actions which, according to these representations, can be expected to achieve that goal (Stewart 2010: 320).

Here it looks likes there is a "purely mental activity" independent of actually acting in the world, i.e., independent of physical behavior in the world. So, the discussion in Stewart, 1996, suggests that one can be an enactivist without being an enactivist$_b$.

6. Conclusion

At the heart of this paper is the observation that some enactivists do not mean by "cognition" what traditionalists have meant by "cognition." There are, if you will, two concepts of cognition in play, a traditional concept and an enactivist concept. This observation would seem to be entirely unproblematic. Moreover, it would seem to be entirely unproblematic to note that some enactivists use "cognition" to describe (a kind of) behavior. These enactivists maintain that cognition is (viable) behavior. These enactivists are enactivists$_b$. This choice of terminology—or this way of theorizing, if you will—however, looks to be misleading. Moreover, it is not misleading just for traditionalists. It is misleading for even some enactivists$_b$. By adopting a new conception of cognition—by thinking of behavior and cognition as the same thing—enactivists$_b$ sometimes overlook ways in which they have detached themselves from the traditions of cognitive science. Enactivists are generally happy to break with these traditions, but there are also times when this break is not as complete as it should be. There are times when they try to engage with mainstream cognitive science, but are hampered by the steps they have already taken to break with tradition. One cannot solve the traditional mind-body problem, if one is not dealing with (something near enough to) the traditional conception of the mind. One might dissolve the problem or abandon the problem, if one rejects the traditional concept, but one cannot solve it. One cannot argue that cognition is embodied and extended, by observing that behavior is embodied or extended. And, one cannot show that not all cognition involves representation by providing instances of behavior that do not involve representation. None of these observations undermines the enactivist$_b$ approach, much less any other enactivist approaches. They merely draw attention to some missteps in the evolution of enactivism$_b$. Perhaps the safest route for enactivists$_b$ is simply to make a clean break with traditional views. Perhaps enactivists$_b$ should walk away from traditional views and leave them to their own devices.

References

Adams, F., and Aizawa, K. 2008. *The Bounds of Cognition*. Malden. MA: Wiley-Blackwell.

Aizawa, K. 2014. What is this cognition that is supposed to be embodied? *Philosophical Psychology*. **DOI:**10.1080/09515089.2013.875280.

Barbaras, R. 2010. Life and exteriority: The problem of metabolism. E. Di Paolo, O. Gapenne and J. Stewart, eds. *Enaction: Toward a new paradigm for cognitive science*: 89-122). Cambridge, MA: MIT Press.

Block, N. 2005. Review of Alva Noë, *Action in Perception*. *Journal of Philosophy, 102*(5): 259-272.

Bourgine, P., and Stewart, J. 2004. Autopoiesis and cognition. *Artificial life, 10*(3), 327-345.

Brooks, R. A. 1997. Intelligence without representation. J. Haugeland, ed. *Mind design II*: 395-420. Cambridge, MA: MIT Press.

Calvo, P., and Keijzer, F. 2008. Cognition in plants. F. Baluska, ed. *Plant-Environment Interactions: From Sensory Plant Biology to Active Plant Behavior* : 247-266. Berlin: Springer Verlag.

Chalmers, D. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3): 200-219.

Chemero, A. 2009. *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Chomsky, N. 1959. A review of Verbal Behavior. *Language, 35*(1): 26-58.

Clark, A. 2010. Coupling, constitution and the cognitive kind: A reply to Adams and Aizawa. R. Menary, ed. *The extended mind*: 81-99. Cambridge, MA: MIT Press.

Clark, A., and Chalmers, D. 1998. The extended mind. *Analysis, 58*(1): 7-19.

Craver, C. 2007. *Explaining the Brain*. Oxford: Oxford University Press.

Froese, T., Gershenson, C., and Rosenblueth, D. 2013. The Dynamically Extended Mind--A Minimal Modeling Case Study. *arXiv preprint arXiv*:1305.1958.

Di Paolo, E., Rohde, M., and De Jaegher, H. 2010. Horizons for the enactive mind: Values, social interaction, and play. E. Di Paolo, O. Gapenne and J. Stewart, eds. *Enaction: Towards a New Paradigm for Cognitive Science*: 33-87.

Haugeland, J. 1998. Mind embodied and embedded. *Having Thought*: 207-237. Cambridge, MA: Harvard University Press.

Hutchins, E. 2010. Enaction, imagination, and insight. J. Stewart, O. Gapenne and E. Di Paolo, eds. *Enaction: Towards a New Paradigm for Cognitive Science*: 425-450. Cambridge, MA: MIT Press.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.

Hutto, D. D., and Myin, E. 2012. *Radicalising Enactivism*. Cambridge, MA: The MIT Press.

Kim, J. 2005. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.

Maturana, H. R., and Varela, F. J. 1980. *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: Springer.

Newell, A., Shaw, J. C., and Simon, H. A. 1958. Elements of a theory of human problem solving. *Psychological Review, 65*(3): 151-66.

Noë, A. 2004. *Action in Perception*. Cambridge, MA: MIT Press.

Rupert, R. 2009. *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.

Shannon, B. 2010. Toward a phenomenological psychology of the conscious. J. Stewart, O. Gapenne and E. Di Paolo, eds. *Enaction: Toward a New Paradigm for Cognitive Science*: 387-424. Cambridge, MA: MIT Press.

Skinner, B.F. 1957. *Verbal Behavior*. New York, NY: Appleton-Century-Crofts.

Stewart, J. 1996. Cognition = life: Implications for higher-level cognition. *Behavioural Processes, 35*(1): 311-326.

Stewart, J. 2010. Foundational issues in enaction as a paradigm for cognitive science: From the origin of life to consciousness and writing. J. Stewart, O. Gapenne and E. Di Paolo, eds. *Enaction: Towards a New Paradigm for Cognitive Science*: 1-31. Cambridge, MA: MIT Press.

Stewart, J, Gapenne, O, & Di Paolo, EA. 2010. Introduction. J. Stewart, O. Gapenne & E. Di Paolo, eds. *Enaction: Toward a New Paradigm for Cognitive Science*: vii-xvii.

Thagard: 2010. Cognitive science. Retrieved November 24, 2013, 2013, from http://www.science.uva.nl/~seop/entries/cognitive-science/

Varela, F., Thompson, E., and Rosch, E. 1991. *The Embodied Mind*. Cambridge, MA: MIT Press.

Webb, B. 1994. Robotic experiments in cricket phonotaxis. D. Cliff: Husbands, M. J. A. and S. W. Wilson, eds. *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*: 45-54. Cambridge, MA: MIT Press.

Webb, B. 1996. A cricket robot. *Scientific American, 275*(6): 94-99.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0003

Enacting anti-representationalism.
The scope and the limits of enactive critiques of representationalism

Pierre Steiner
Université de Technologie de Compiègne – Sorbonne Universités
COSTECH – Cognitive Research and Enaction Design
Pierre.Steiner[]utc.fr

Received August 2014; accepted September 2014; published Autumn 2014.

Abstract

I propose a systematic survey of the various attitudes proponents of enaction (or enactivism) entertained or are entertaining towards representationalism and towards the use of the concept "mental representation" in cognitive science. For the sake of clarity, a set of distinctions between different varieties of representationalism and anti-representationalism are presented. I also recapitulate and discuss some anti-representationalist trends and strategies one can find the enactive literature, before focusing on some possible limitations of eliminativist versions of enactive anti-representationalism. These limitations are here taken as opportunities for reflecting on the fate of enactivism in its relations with representationalism and anti-representationalism.

Keywords: natural content; mental representation; representationalism; enactivism; anti-representationalism; theoretical terms; eliminativism.

Introduction

The criticism and the rejection of representationalism have a particular status in enactivism (or enaction). The definition of cognition as *embodied action* was explicitly proposed by Varela, Thompson and Rosch in their seminal *The Embodied Mind* as an alternative to the definition of cognition as "the representation of a world that is independent of our perceptual and cognitive capacities by a cognitive system that exists independent of the world" (1991: xx). Still, if we are looking nowadays for a more positive definition of enaction, concepts such as *autonomy*, *autopoiesis*, *embodiment*, *structural coupling*, *sense-making*, *life-mind continuity*, or *lived experience* immediately and eminently come to the fore, somehow leaving the criticism of representationalism in the background[17]. It is true that the unpacking of the meanings of these concepts often carries or implies an implicit criticism of representationalism as the one mentioned above, and is very helpful for contrasting enactivism with other non-representationalist theories of cognition such as radical embodied cognitive science (Chemero 2009). Nevertheless, enactivists would be wrong to consider that developing an explicit criticism of representationalism was only necessary when enaction emerged as an alternative theory (or even paradigm) to cognitivism, where the concept of "mental representation" was—and is still—a basic building block. Strategically, the rejection of representationalism continues to mark an important difference not only between enactivism and cognitivism (i.e. the computo-representational theory of mind), but also between (some forms of) enactivism and other more recent theories that criticize, amend or even reject the intellectualist, internalist or formalist dimensions of the computo-representational theory of mind… by often retaining representationalism. Amongst these recent theories, one can include distributed cognition (Hutchins 1995), situated cognition (Clancey 1997), extended cognition (Clark 2008), and interactivism (Bickhard 2009)[18]. More fundamentally, the permanent rejection of representationalism should not be taken as an *easy* task: representationalism is a polymorphous and plastic thesis, sometimes looking like a Lernaean Hydra: shallow and maximalist versions of representationalism are easy to see and to dislodge, but only at the benefit of other versions which are much more tenacious, refined and pervasive. It is the implicit endorsement of these latter versions that may explain why some authors propose "enactive" accounts of mental representations (Ellis and Newton 2010), or that some reviewers of the enactivist literature state that "nothing in the enactivist view requires abandonment of contentful states" (Shapiro 2014).

Let us take some examples, by contrasting three different understandings of what anti-representationalism may amount to:

> (1) Cognition and cognitive phenomena such as perception, language-understanding or problem-solving are not representational, in the sense of being functionally isolated from action or from the active and embodied engagement of cognitive creatures in the world;
>
> (2) Cognition and cognitive phenomena such as perception, language-understanding or problem-solving are not representational, in the sense of consisting in the manufacture, the manipulation or the retrieval of

[17] See for instance the contributions in the book edited by Stewart, Gapenne and Di Paolo 2010.

[18] In a recent paper (Steiner 2014), I have even argued that the extension of cognition in the world (and not only in bodily engagements with the world) is very restrained if representationalism (even minimal) is retained.

symbolic, abstract, action-neutral and detailed mental representations of the environment;

(3) Cognition and cognitive phenomena such as perception, language-understanding or problem-solving are not representational, in the sense of involving the manufacture, the manipulation or the retrieval of mental representations as contentful physical structures (be they intracranial or distributed across brain, body and world), whatever their formats and roles in cognitive processing.

(1) concerns the *role* of representation in the definition of cognitive phenomena, (2) concerns the *formats* and the *properties* of mental representations, whereas (3) is about the very existence of mental representations (whatever their roles, properties and formats). Endorsing (1) and (2) is not sufficient for endorsing (3). Indeed, most if not all enactivists will clearly endorse (1) and (2), but they will not be alone doing so: many friends of 3E-embodied, embedded, extended-cognition, including situated cognition and distributed cognition will also endorse (1) and (2). One might expect that the difference is or at least should be made in the endorsement of (3). Still, most if not all enactivists may find (3) to be too brutal or radical, since (3) denies that there are *any* mental representations involved in cognitive processes. For instance, after making it clear that he argues “against representationalist theories that separate perception and action (…) and that neglect the ways autonomous agents bring forth or enact meaning in perception and action” Evan Thompson (2011: 194) expresses his sympathy for mental representations as they are defined in the emulation theory of mental imagery (Foglia & Grush 2011). Another example can be found in O’Regan and Noë’s “A sensorimotor account of vision and visual consciousness”: the authors reject the claim that vision requires the production or use of detailed representations, but they still accept that the visual system stores and uses information, and that “seeing lies in the *making use* of the representation, not in the having of the representation” (2001: 1017; their emphasis). Finally, (3) might be associated with what Dan Hutto and Erik Myin (2013) name “really radical enactivism”, a position they do *not* claim to endorse (2013: xviii). For “really radical enactivism”, cognition *never* involves representational content. The basic claim of Hutto and Myin’s radical enactivism is rather that only basic cognition (typically exemplified in perceptual experience, sensori-motor coordination, reaching and grasping, or keeping track of another’s gaze) is not contentful or representational, even if it exhibits intentional directedness. Their non-endorsement of (3) is thus different from the non-endorsement of (3) as we can find it in Noë and Thompson: Hutto and Myin endorse (3) for what they call “basic cognition”, but not for other forms of cognition (such as cases of linguistic judgments or intelligent planning (2013: 40-41)); whereas Noë and Thompson do not claim that basic cognition is non-representational. Radical enactivism admits that enculturated or

linguistically-scaffolded minds may be informed by or involve contents or mental representations (Hutto and Myin 2013: ix: xviii: 82).

Consider representationalism as being the existential claim that *there are mental representations as contentful physical structures playing a role in cognitive processing*. One might think that the radicality of *radical enactivism* consists in the fact it rejects representationalism as applied to basic cognition (non-radical versions of enactivism, like Thompson or Noë, do not do that). But this characterization of the radicality of radical enactivism presupposes that a clear line could be drawn between basic cognition and other kinds of cognition, and that this line parallels the "non-representational/representational" distinction. In the case of human cognition at least, it is questionable that basic cognition does not involve representational content: if basic cognition is acquired and exercised in socio-cultural practices, there are good reasons to think—if we follow Hutto and Myin—that it is contentful, so that the non-representational dimensions of basic cognition would be very marginal. Because of this possible difficulty of the distinction between basic cognition and non-basic cognition, I think it is preferable to underline and to define the radicality of radical enactivism alternatively.

If representationalism is very basically defined as the existential claim that *there are mental representations as contentful physical structures playing a role in cognitive processing*, the following table can be helpful for summarizing the current situation:

Mental representations as… / Theory	Existing as contentful physical structures playing a role in cognitive processing	Necessarily symbolic, detailed, abstract	Necessarily intracranial	Necessarily involved in all cases of cognitive processing	Not involved in cases of basic cognition	Possibly made out of natural content	Necessarily made out of linguistic and social resources
Computo representational theory of mind	Yes	Yes	Yes	Yes	No	Yes (necessarily made out of neural content)	No
Parallel Distributed Processing	Yes	No	Yes	Yes	No	Yes (necessarily made out of neural content)	No

Extended Cognition, distributed cognition	Yes	No	No	Yes/No (depending on the author)	No	Yes (but not necessarily made out of neural content)	No
Enactivism (Thompson, Noë)	Yes	No	No	No	***No***	***Yes*** (but not (necessarily made out of neural content)	***No***
Radical Enactivism (Hutto & Myin)	***Yes***	No	No	No	***Yes***	***No***	***Yes***
Really Radical Enactivism	***No***						

Table 1

By looking at this chart, one may note that radical enactivism is rejecting a very basic kind of representationalism which is logically narrower than the representationalism just defined above, but which also constitutes the core assumption of the great majority of actual versions of representationalism: the existential claim that there are mental representations, defined as physical structures (vehicles) playing a role in cognitive processing *in virtue of some content whose existence does not depend on the existence of social and linguistic practices and on the ability of the organism to take part in those practices*[19]. Typically, these mental representations consist in subpersonal and intracranial processes that are naturally or intrinsically contentful[20]. Radical enactivism refuses that idea by arguing that

> contents and vehicles exist, but they are associated with linguistic symbols and forms of cognition that feature in and are logically and developmentally dependent upon shared, scaffolded practices (Hutto and Myin 2013: 152).

[19] According to this definition, the claim that cognitive processing involves the use of public representational systems or the production of personal-level representing mental acts that consist in the internalization of public representations is *not* a representationalist claim.

[20] A synonym for "made out of natural content" is "intrinsically having content": the possession of content by physical vehicles does not depend on the existence of linguistic, representational or symbolic human practices. "Intrinsically" does *not* mean here "non-relational".

To put it otherwise: radical enactivism claims that cognition never involves mental representations made out of natural content. This is an idea that is radical enough for serving here as a starting point for defining the clearest current form of enactive anti-representationalism. Radical enactivism is now sufficiently radical when it is compared with classical internalist versions of representationalism for which mental representation necessarily have intracranial vehicles carrying natural content, and with "extended mind" and other enactive versions of representationalism for which there may be mental representations (intracranially located or not) made out of natural content (besides other types of mental representations, including public-language representations and external representations). If we define representationalism as a claim being about representations endowed with natural or intrinsic content, the radicality of radical enactivism can appear in the following table:

Mental representations as… / Theory	Existing as being made out of natural content	Necessarily symbolic, detailed, abstract	Necessarily intracranial	Necessarily involved in all cases of cognitive processing
Computo representational theory of mind	Yes	Yes	Yes	Yes
Parallel Distributed Processing	Yes	No	Yes	Yes
Extended Cognition/distributed cognition	Yes (but there are also other types of mental representations)	No	No	Yes/No (depending on the author)
Enactivism (Thompson, Noë)	Yes (but there are also other types of mental representations)	No	No	No
Radical Enactivism (Hutto & Myin)	***No***			

Table 2

In this paper, I wish to clarify and to assess some arguments proponents of enactivism (radical or not) have proposed in their criticism of representationalism (in the narrow sense just defined above). This will first require a set of distinctions about the *targets* and the *forms* of enactive anti-representationalism (section I). I will then proceed by rehearsing two *classical* strategies against representationalism that the enactivist tradition has exploited but that can also be found in other traditions (section II), before presenting two more *radical* and *specific* anti-representationalist strategies that I see as being more proper to the enactivist tradition (section III). Still, these two latter strategies might pose no problems for a marginal yet existing version of representationalism, according to which mental representations do not essentially have content and intentionality (section IV). This resistance of a marginal version of representationalism to the (proclaimed) radicality of enactive anti-representationalism will absolutely not be considered here as a refutation or dismissal of the latter. On the contrary, it will be seen as an opportunity—in section V—to identify some common assumption(s) that enactive anti-representationalism and classical representationalism might share, but also to invite enactivist anti-representationalists to (re)consider the conditions by which theoretical terms may be eliminated (or retained) in science.

Allow me to end this introductory section with a personal note: I am writing here from a (global) anti-representationalist stance (defended in other papers)[21]; I will definitely not present here a critique of representationalism, or a defense of anti-representationalism. My aim here is to describe the scene from which enactive anti-representationalism has been and is currently enacted. This description is a requisite for a better understanding, refinement, but also possible criticism of enactive anti-representationalism.

I. Situating anti-representationalism

In order to precisely define the various forms of enactive anti-representationalism, it is necessary to situate them among a broader *Spielraum* defined by at least twenty-four possible positions (don't worry: only twelve of them will be actually considered!). These twenty-four positions are constructed out of the combinations between the choices that can be made when one is facing three main alternatives: an alternative between *positions* (representationalism vs. non-(or anti-) representationalism); an alternative between the *stances* from which these positions are defended (methodological vs. ontological); and an alternative between the *scopes* of these positions (local vs. basic vs. global). I have presented above preliminary definitions of representationalism and anti-representationalism: it is now time to sharpen them. But let me remind you that all these clarifications and distinctions will be made from a common starting point: representationalism as a claim about the existence of physical structures endowed with *natural* or *intrinsic* content and playing a role in cognitive processing.

[21] See Steiner (2010) for a critique of representationalism and a defense of anti-representationalism; Steiner (2014) for a critique of extended-mind representationalism; and Steiner (2011) for a defense of enactive anti-representationalism.

A very basic distinction is classically drawn between *methodological* and *ontological* versions of representationalism and anti-representationalism[22]. Ontological versions explicitly take issue with the reality of mental representations. Ontological representationalism considers that mental representations exist, whereas ontological anti-representationalism denies their existence. A methodological version is agnostic on the issue of the reality of mental representations: they are not taken as literally existing (we may not be warranted in positively asserting their reality). Still, methodological representationalism argues that mental representations (and their manipulation, manufacture or retrieval) must necessarily be posited for the explanation or prediction of the performances of cognitive systems; whereas methodological non-representationalism argues that they do not need to be invoked (it is possible to posit them, but there are more helpful theoretical posits). In both methodological stances, nothing ontological is inferred from the presence or the absence of the concept "mental representation" in successful explanatory and predictive practices (methodological representationalism includes "fictionalism" about mental representations (Sprevak, 2013)). Methodological versions of representationalism and anti-representationalism may be said to be anti-realist in the following sense: they deny that theories involving the positing (or the non-positing) of mental representations are truth-conditioned descriptions of their intended domain (observable and unobservable), and that their predictive or explanatory successes entail that the entities they posit have "real" counterparts.

Methodological *non*-representationalism is not a variety of *anti*-representationalism, since it basically makes *no* use of the concept "mental representation". In itself, it is not *against* the existence of mental representations: it may consider that the ontological debate between representationalism and anti-representationalism is vain, for instance because of a lack of clear definition of what a representational property is[23]. Anti-representa-tionalism is more demanding and challenging than non-representationalism, since it explicitly claims that mental representations do not exist.

The endorsement of an ontological position does not force one to make a choice concerning a particular methodological commitment: both ontological representationalism and anti-representationalism are compatible with both methodological representationalism and non-representationalism, but also with the attitude of having no commitment at all towards the methodology of cognitive science. And conversely: methodological commitments may be independent of ontological commitments and interests. Still, one may combine

[22] See Chemero (2000) for this distinction (but I use the term "methodological" where Chemero used "epistemological").

[23] See Haselager et alii. (2003).

an ontological position with a methodological position, and thus endorse at the same time both an ontological *and* a methodological position.

We thus have eight possible positions:

– Ontological representationalism;

– Methodological representationalism (which I will denote hereafter by “methodological representationalism”);

– Ontological representationalism with methodological representationalism (which I will denote hereafter by “representationalism”);

– Ontological anti-representationalism;

– Methodological non-representationalism (which I will denote hereafter by “methodological representationalism”);

– Ontological anti-representationalism with methodological non-representationalism (which I will denote hereafter by “anti-representationalism”);

– Ontological representationalism with methodological non-representationalism: one accepts that mental representations exist, but consider that scientific models should better do without the concept of “mental representation”.

– Ontological anti-representationalism with methodological representationalism: one accepts that mental representations do not exist, but still holds that they are our best ways to capture and explain the complexity of cognitive behaviour.

In ontological representationalism and in ontological anti-representationalism, one does not want to infer methodological consequences from the ontological position, or does not want to ground this ontological position on methodological commitments. In methodological representationalism and methodological non-representationalism, one defends a claim concerning the methodology of cognitive science, but does not want this methodological choice to interfere with ontological issues.

The scope of each of these eight positions can be global, basic or local.

The position is *global* when it applies to *every* cognitive system and *every* cognitive part of it (operations and subsystems such as faculties). It is *basic* when it applies to *most* cognitive systems, operations and subsystems. And it is *local* when it only applies to the particular cognitive system, operations or subsystem under consideration.

We thus have twenty-four (8 x 3) different positions. I will focus on what I consider as being the twelve most notable positions. Indeed, for obvious reasons of space but also of relevance, I leave aside global, basic and local versions of ontological representationalism and anti-representationalism: in the philosophy of cognitive science, there are not many scholars who defend ontological representationalism or anti-representationalism *only*, not aggregating them with methodological commitments or suggestions. Ontological representationalism and anti-representationalism are almost always included as components of what I call here "representationalism" and "anti-representationalism", which also include methodological commitments. I also leave aside global, basic and local versions of conjunctions of ontological representationalism with methodological non-representationalism, and of ontological anti-representationalism with methodological representationalism. These versions are quite rare in the literature, and it is hoped that the reader will be able to define them from the statements given above. This leaves us with 12 positions, which we will now examine one-by-one.

(1) *Global representationalism*: Every cognitive system and every cognitive part of it (operations and subsystems such as faculties) involves the use, the retrieval or the manufacture of mental representations (as made out of natural content), so that *bona fide* models of every system, operations or subsystem as *cognitive* system, operations or subsystem must appeal to the concept of "mental representation"—for descriptive, predictive and explanatory purposes.

(2) *Basic representationalism*: Most cognitive systems, operations and subsystems include the use, the retrieval or the manufacture of mental representations (as made out of natural content), but there might be cognitive systems, sub-systems and operations that do not include mental representations (made out of natural content or not): models of these systems, operations or sub-systems as *cognitive* systems, operations or subsystems do not need to appeal to the concept of "mental representation" for descriptive, predictive and explanatory purposes.

(3) *Local representationalism*: the cognitive system, operations or subsystem under consideration includes the use, the retrieval or the manufacture of mental representations (as made out of natural content), so that *bona fide* models of this system, operations or sub-system as *cognitive* system, operations or subsystem must appeal to the concept of "mental representation" for descriptive, predictive and explanatory purposes.

All these cases of representationalism are composed of ontological representationalism with methodological representationalism. Mental representations are here contentful (information-carrying) physical structures that have a real ontological and explanatory status. Their content is a natural product that allows them to refer to some object, property or state of affairs. According to this definition, external (public and/or shareable) representations such as models, images or natural language sentences are *not* mental representations: one can be a critic of representationalism without denying the existence and the cognitive importance of external representations. Representationalism is here a claim about the *reality* of the property "being a mental representation of X": it is not only about its heuristic relevance for describing and explaining the mechanisms of cognition. Representationalism, as we define it here, is a very general claim, that includes many variations: one can find a version of representationalism for which all mental representations are necessarily symbols in a language of thought (Fodor, 1987, chap.1 and appendix), or a version of representationalism for which cognition must be defined as a set of operations having the function of building mental representations of environmental phenomena (in that version, the property of mental representation is used for defining the *explanandum*, and not only the *explanantia* of cognitive science). These two versions of representationalism can be criticized, amended or even rejected by other versions of representationalism (for instance: parallel and sub-symbolic distributed processing vs. the symbolic conception; action-oriented conceptions of cognition and representation vs. the idea of cognition as a mirror of the environment ; mental representations as maps, models or pictures vs. mental representations as propositional sets of symbols…). Global representationalism does not necessarily link the cognitive character of a system to the presence of mental representations: it just asserts that from some level of study and analysis of cognitive systems, it is necessary to acknowledge the existence of mental representations, without assuming that they correspond perfectly to what would be described at another level of analysis of these same systems. Still, global representationalism considers that explaining the cognitive properties of cognitive systems requires the appeal to mental representations.

Vehicle-internalist and cognitivist theories of cognition are not the exclusive owners of representationalism. On the contrary: many versions of extended, distributed or situated cognition may endorse basic representationalism. For most proponents of extended cognition, even if there may be cases or aspects of extended cognitive processing that do not (just) involve mental representations as made out of natural content and thus that do not require representationalist explanations, the existence of mental representations (for instance realized in intracranial and subpersonal processes) and the necessity of referring to them when one tries to explain a great variety of cognitive phenomena are not at issue[24]. Unlike classical AI representations, these representations do not need to be complete, inert, propositional, denotational, action- and perception-neutral, stable, complex, detailed, digital, discrete, amodal, syntactically structured, or symbolic. Mental representations can be built and used on the fly; they can be modal (even when they are categorical), minimal (content-sparse), superposed, partial, action-oriented, context-dependent, embodied, analogue, distributed, or sub-symbolic. Basic representationalism is more

[24] See for instance Clark (2008: 19, 26, &153); Menary (2007: 58–59 & 69); Rowlands (2010, chapter 5); Sutton (2010: 197); Wheeler (2005, chapter 8); and Wilson (2004, chapter 8).

plastic and flexible than global representationalism: it easily tolerates the existence of cognitive phenomena that are not representational (and that do not deserve representational explanation). While position (1) is rather endorsed by standard versions of the cognitivist theory, post-cognitivist theories that insist on the pragmatic, environmental, embodied or situated dimensions of cognition by retaining representationalism generally endorse basic representationalism.

(4) *Global methodological representationalism*: *Bona fide* models of every system, operation and subsystem as cognitive systems, operations or subsystems must appeal to the concept of "mental representation" for predictive or explanatory purposes.

(5) *Basic methodological representationalism*: *Bona fide* models of most systems, operations and subsystems as cognitive systems, operations and subsystems must appeal to the concept of "mental representation" for predictive or explanatory purposes, but there might be systems, operations and subsystems whose cognitive properties can be explained or predicted without appealing to the concept "mental representation".

(6) *Local methodological representationalism*: *Bona fide* models of this system, operation or sub-system as *cognitive* system, operation or subsystem must appeal to the concept "mental representation" for predictive or explanatory purposes.

It is important to note that methodological representationalism may accept that there might be non-representationalist descriptions of cognitive systems: it will just argue that, when it comes to explanation and to (interesting) predictions, the ascription of mental representations (as made out of natural content) is the only possible way (globally, basically or locally). Methodological representationalism is well exemplified in the daily practices of many researchers in cognitive science: the use of the concept "mental representation" is considered as being absolutely required for describing, predicting and explaining studied phenomena as *cognitive* phenomena, but nobody will dare to enter into ontological considerations by asserting that *mental representations* exist (or not). That is, the endorsement of methodological representationalism carries no commitment to the existence or non-existence of mental representations.

I mentioned above that methodological versions of representationalism (and anti-representationalism) were anti-realist towards the property "being a mental representation". There are different kinds of anti-realism: instrumentalism, but also pragmatism, phenomenalism, interpretationism, constructive empiricism, fictionalism and idealism, to name but a few. The combination of methodological representationalism with each of these versions of anti-realism can foster complex and subtle versions of methodological representationalism. For some of them, methodological representationalism is necessary (this is the version presented here), for other versions, it is only a better position than non-representationalism, and for other versions, both representationalism and non-representationalism are possible (and efficient) ways of describing, explaining or predicting the behaviour of cognitive systems. Depending on the kind of anti-realist commitments one assumes when methodological representationalism is endorsed, mental representations can have different status: they can be conceived as models, fictions, useful falsehoods, explanatory tools, instruments of calculation, descriptive labels,… In any case, if the concept "mental representation" is a representation, it is not a representation of an object "out there", but a part of a representational system such as a theory, whose main purpose is not describe or to represent what there is in an unobservable domain, but rather to predict and/or to explain the behaviour of cognitive systems. Mental representations are not constituents of cognitive systems; they are constituted in and by scientific practices and theories. The fact one posits mental representations—and not nails, flies or cucumbers—for achieving predicting and explanatory purposes can be explained—but not justified—by mentioning how it is often comforting to rely on representational systems such as language for modeling and defining thought and its intentionality (Sellars 1956, § 50-52, and § 57-58 is a central reference on this topic).

(7) *Global non-representationalism*: *Bona fide* models of every system, operation and subsystem as *cognitive* system, operation and subsystem do not need to appeal to mental representations (as made out of natural content), for descriptive, predictive and explanatory purposes.

(8) *Basic non-representationalism*: *Bona fide* models of most systems, operations and subsystems as *cognitive* systems, operations and subsystems do not need to appeal to mental representations (as made out of natural content), for descriptive, predictive and explanatory purposes, but there might be systems, operations or subsystems whose cognitive properties need to be described, explained, or predicted by appealing to mental representations.

(9) *Local non-representationalism*: *Bona fide* models of this system, operation or sub-system as *cognitive* system, operation or subsystem do not need to appeal to the concept "mental representation" for descriptive, predictive and explanatory purposes.

The three cases of non-representationalism are not symmetrical with the three cases of methodological representationalism, since methodological representationalisms express a necessity in the form of a normative claim ("one *must* appeal to the concept of 'mental representation'") while non-representationalisms deny this necessity without implying that non-representationalist models and explanations should automatically be *preferred* to representationalist models and explanations. And remember that non-representationalism do not claim that representationalist explanations are *mistaken*. It just holds that they can be dispensed with (globally, basically or locally).

(10) *Global anti-representationalism*: There are no mental representations (understood as physical structures having natural content), so that (a) cognitive systems, operations and sub-systems do not include the use, the retrieval or manufacture of mental representations and (b) *bona fide* models of systems, operations and subsystems as cognitive should not appeal to mental representations (and thus use the concept "mental representation") for descriptive, predictive and explanatory purposes.

(11) *Basic anti-representationalism*. Most cognitive systems, operations and subsystems do not include the use, retrieval or manufacture of mental representations, so that *bona fide* models of these systems, operations and subsystems as cognitive should not appeal to mental representations (and thus use the concept "mental representation") for descriptive, predictive and explanatory purposes.

(12) *Local anti-representationalism*: the cognitive system, operations or subsystem under consideration does not include the use, the retrieval or the manufacture of mental representations (as physical structures having natural content), so that *bona fide* models of this system, operations or sub-system as *cognitive* system, operations or subsystem should not appeal to mental representations (and so should not use the concept "mental representation") for descriptive, predictive and explanatory purposes.

Since representationalism and anti-representationalism result from a conjunction of ontological and methodological positions, each of these latter positions can *partially* support representationalism and anti-representationalism Global representationalism and global anti-representationalism are more than methodological positions, and concern every cognitive system: empirical statements and examples will not be sufficient for justifying them. Moreover, global, basic and local versions of anti-representationalism include ontological claims on the non-reality of mental representations. These ontological claims will be conceptual, not empirical, for one cannot ask to the proponent of anti-representationalism to empirically show that mental representations do not exist[25]. But anti-representationalisms also include methodological components: arguments related to the possibility of non-representational explanatory practices in cognitive science can partially justify them.

[25] Indeed, an existential claim such as "There are mental representations" could only be falsified by a negative existential claim such as "There are no mental representations". But a negative existential claim is a *universal* claim, and these claims cannot be *empirically* confirmed (see Popper 1959, chap. 3, section 15 for that classical point).

Ontological representationalism is included in representationalism: a local version of ontological representationalism (or a local version of representationalism) can be used for refuting global versions of ontological anti-representationalism or global versions of anti-representationalism. Indeed, the simple observation of a mental representation is sufficient to refute the claim that they do not exist. But in order for the observation of a mental representation to refute global ontological anti-representationalism, it is first of all necessary to define the necessary and sufficient conditions for the presence of a mental representation, which is not at all obvious if one takes into account the various debates inside of representationalism itself.

Any local version of representationalism is compatible with local or basic versions of anti-representationalism (or non-representationalism), and any local version of anti-representationalism is compatible with local or basic versions of representationalism (ontological and methodological, or methodological only). Those who endorse these aggregated positions will often hold that representationalism and anti-representationalism can be, or even have to be, *complementary* approaches. It is only for global versions that representationalism and anti-representationalism are *contradictory* approaches.

Now that these (hopefully) clarifying distinctions have been made, we can ask: what kind of anti-representationalism can we find in the enactive literature?

First, let us recall that none of these positions can suffice for *defining* enactivism, since enactivism is not only a claim about the representational (or non-representational) properties of cognitive systems. Answering two other questions may help in answering the question raised above.

> (A) Among the six different versions of representationalism (1, 2, 3, 4, 5, and 6), which version(s) does enactivism reject?
>
> (B) Among the six different versions of anti-representationalism (7, 8, 9, 10, 11, and 12), which version(s) does enactivism endorse?

Of course, the answer given to (A) will depend on the answer given to (B), and conversely.

One can identify at least three broad trends in the current enactive literature concerning the ontological and the explanatory status of mental representations:

The first trend is an explicitly *eliminativist* trend, consisting in the explicit defence of global anti-representationalism, and thus in the complete rejection of all the six versions of representationalism, ontological and methodological. Hutto and Myin's *radical enactivism* is here a landmark. The battle against representationalism is here engaged on two fronts: an ontological front (mental representations as made out of natural content do not exist), and a methodological front (non-representationalist explanations are possible and should be favored). But we must keep in mind that radical enactivism is *not* "really radical enactivism": radical enactivism accepts that some classes of cognitive phenomena exhibit representational properties, but these representational properties necessarily depend on the existence of sociocultural practices (representational properties are not natural properties). Since radical enactivism denies the very existence of natural content, it is coherent here to define it as defending a global version of anti-representationalism (as we have defined it).

The second trend is a *conservative* trend, rejecting (1), (2), (4) and (5) by endorsing (8) or (11). There is thus non-representationalism or even anti-representationalism here, but conceding that some forms of mental representations (as made out of natural content, and possibly realized in subpersonal and intracranial processes) may exist and/or that representationalist explanations may be required in some cases of cognition, like for instance anticipation, abstraction, imagination, or memory. Conservative enactivism is thus compatible with (3) and (6). Conservative enactivists will here insist that these mental representations are not symbolic, static, abstract or detailed, but they will endorse representationalism. I have mentioned Evan Thompson's and Alva Noë's acceptances of some forms of representationalism in the introductory section. John Stewart also endorses a version of what I call here "conservative trend" when he suggests that

> "Constructivist representations", if I may call them that, cannot of course represent referential states of affairs in the external objective world (as in computationalism). I consider, however, that they can (…) represent the anticipated consequences of an organism's actions for its future perceptions. Armed with representations of this sort, an organism can set itself a "goal" (expressed in terms of a desired perceptual configuration), and then by purely mental activity (without having to take the risks involved in proceeding by trial and error by actually acting in the world) elaborate a sequence of actions which, according to these representations, can be expected to achieve that goal (…). Representations as thus defined are thus the basis for intentional action. (Stewart 1996, III.5).

The author ably criticizes and rejects mental representations as they figure in a computational and objectivist theory of cognition, but accepts that the enactive theory (equated with a constructivist approach) can harbour other types of mental representations (presumably endowed with natural content, since the author does not mention some necessary dependence between mental representations and socio-cultural practices) . Since it is quite easy to find versions of representationalism that reject or are agnostic on the metaphysical debate between objectivism and constructivism (Clark 1997: 173) and/or that do not endorse the computational theory of mind, one can understand why this conservation of some types of mental representations *is* a type of (local) representationalism.

The conservative trend faces at least two dangers: first, the proponents of basic representationalism (be it ontological and methodological or methodological only) can argue that the cognitive phenomena "conservative enactivists" agree to define or to model in a representationalist way are actually so pervasive or fundamental in our cognitive life that the enactivist should accept, by implication, that our cognitive life *is* basically representational (in Stewart's case: are not intentional actions pervasive in our cognitive life?). Conservative enactivism will then slowly but surely slide from local representationalism (compatible with basic anti-representationalism), be it (9) or (12), *to* basic representationalism (which is incompatible with basic anti-representationalism), be it (8) or (11). Second, if they are able to stick to local representationalism and argue that they endorse basic anti-represen-tationalism, conservative enactivists need to explain why they endorse basic anti-representationalism. Eliminativist arguments on any kind of mental representations as they are developed by or for global anti-representationalism (ontological and methodological, or methodological only) will not be available to them, since conservative enactivists agree to see that some mental representations (as having natural content) exist (or should be taken as existing, for explanatory purposes). Conservative enactivists will need to carefully decompose the architecture of cognitive functions in order to state where and how mental representations exist (or should be posited as existing) and where and how they do not exist. This option for criticizing representationalism is possible, of course, but is trickier and less encompassing than the roads which are taken by global anti-representationalism and non-representationalism.

But the very possibility of the conservative trend (that is, the fact that some enactivists are ready to embrace it) may also reveal a shortcoming of the eliminativist trend: the need of retaining a representationalist vocabulary for explaining anticipation (for example) may be due to the current non-availability of other (non-representational) concepts in the toolbox of global anti-representationalism or non-representationalism. Or, alternatively, the proponent of the conservative trend may consider that the non-representational concepts anti-representationalists or non-representationalists want to introduce in place of representational concepts are currently not adequate for describing or explaining the cognitive phenomenon under question, and notably their intracranial and subpersonal components.

The last trend in the enactive literature is a *practically eliminativist* trend, endorsing global methodological non-representationalism without embracing global anti-representationalism: the aim of this trend here is to move forward by developing applications of non-representationalism, leaving ontological controversies on mental representations behind. The proponent of the agnostic trend does not *claim* that mental representations or the representationalist language should be eliminated in enactive cognitive science; he may just be indifferent to the issue of representationalism. Spending time criticizing representationalism, as an explanatory commitment or as an ontological stance, would already be giving too much importance to classical cognitive science.

According to this trend, it by its practical fruits that non-representationalism must be considered and defended, not by the criticism of representationalism. An example of such an attitude can be found in Chemero's (non-enactive) non-representationalism (for which non-representationalist explanations are not only possible, but also better than representationalist explanations):

"Refrain from arguing that cognitive systems *really* are not representational; instead, argue that the best way to understand cognition is with the tools of dynamical systems theory, by taking up what I have called the dynamical stance. The best way to argue for the fruitfulness of the dynamical stance is by example; get to work providing non-representational explanations of cognitive phenomena that are both convincing and sufficiently rich in their implications to guide further research" (2000: 646; author's emphasis).

Replace here "dynamical systems theory" by "autopoietic enactivism" and "dynamical stance" by "enactive stance" and you get a practically eliminativist trend in enactivism (of course, the resources of dynamical systems theory may be integrated in enactivism, and conversely). Note that this option is only a distinct option if one claims (or believes) that non-representationalism is *globally* possible. A restriction of its scope would lead to the conservative trend (for instance, to position 8).

To sum up: the *absence of the concept "mental representation" in an enactive theory does not necessarily entail global anti-representationalism*. Every time a proponent of enactivism develops a model of some cognitive phenomenon without using a representationalist vocabulary, he might be doing so for different reasons or from different hypothesis: he may endorse the eliminativist trend of anti-representationalism, and thus global anti-representationalism; he may endorse basic anti-representationalism, and thus be disposed to endorse the representationalist idiom in order to account for some cognitive phenomena; or he may be indifferent to the representationalism vs. anti-representationalism debate, and rather committed to the attempt of accounting for cognitive phenomena in general with a non-representationalist vocabulary, at best by *implicitly* endorsing a global non-representationalist commitment (option 3). The proponent of the practically eliminativist trend might refuse to make explicit his ontological commitments on the issue of mental representations… because he might have no commitments on that topic! In this latter sense, it is possible that the proponent of the practically eliminativist line has no desire to build a systematic anti-representationalist theory: non-representationalism in action is enough.

The existence of these three trends *within* enactivism is both normal and problematic. It is normal, for it would be naive (and even dangerous) to expect too much homogeneity from enactivism. There are different ways of embracing the computational theory of mind or distributed cognition: why would that be different for enactivism? But it is also problematic, for these three different trends exemplify different potential (and often current) tensions within enactivism, which may weaken its unified criticism of other theories. For instance, the conservative trend implicitly disagrees with the explicitly eliminativist trend: it accepts neither the scope of its conclusion nor its main ontological claim… for this ontological claim rules out the very existence of the mental representations the conservative trend wants to retain. The practically eliminativist trend might object that the explicitly eliminativist trend is losing time by explicitly criticizing a lost cause (ontological representationalism): it should only focus on practical work (the practical development of methodological non-representationalism). The conservative trend can also accuse the practically eliminativist trend of not providing satisfying alternatives to representationalist tools, but it can also be accused—by eliminativists—of conceding too much to representationalism. All of this means that one of the next important challenges of enactivism is not the criticism or the rejection of representationalism *per se*, but the clarification of how and why representationalism should be criticized, abandoned… or retained (let us also not forget that "really radical enactivism" can also be a contender in the debate !).

I have said above that it would be naive (and even dangerous) to expect too much homogeneity from enactivism. Besides, the various arguments put forward by Varela, Thompson and Rosch in *The Embodied Mind* against representationalism were already difficult to classify. The book certainly included a rejection of the symbolic model of mental representation *and* a rejection of the idea that the concept of "mental representation" is and should be fundamental in cognitive science (see for instance 1991: 9). But this seems to be a version of (11), namely basic anti-representationalism. True, in enactive cognition, "representations no longer play a central role" (1991: 207)… but, therefore—and by pure deduction—they still play a role! Still, in other places, a rejection of any version (symbolic, connectionist,…) of representationalism is expressed, so that representation is not only non-fundamental: it is to be refused. For instance, after having written that they accept the mundane sense of "represent" as expressing a referential property external items (sentences, maps,…) have, the authors criticize a stronger sense of "represent" as it is used for characterizing cognitive activity:

> This strong sense arises when we generalize on the basis of the weaker idea to construct a full-fledged theory of how perception, language, or cognition in general must work. The ontological and epistemological commitments are basically twofold: We assume that the world is pregiven, that its features can be specified prior to any cognitive activity. Then to explain the relation between this cognitive activity and a pregiven world, we hypothesize the existence of mental representations inside the cognitive system (whether these be images, symbols, or subsymbolic patterns of activity distributed across a network does not matter for the moment). We then have a full-fledged theory that says (1) the world is pregiven; (2) our cognition is of this world-even if only to a partial extent, and

> (3) the way in which we cognize this pregiven world is to represent its features and then act on the basis of these representations. (1991: 135)

The authors, here, do not criticize global representationalism (and the five other versions of representationalism) in general: they criticize global representationalism (and the five other versions of representationalism) *as embedded in some objectivist ontology* (see their clauses (1) and (2)). But it is perfectly possible to endorse representationalism without endorsing an objectivist ontology (that is, without endorsing these clauses (1) and (2)). It is not uncommon to read proponents of representationalism arguing that their representationalism, as a hypothesis on the causal and subpersonal mechanicisms of cognition, is totally neutral or agnostic in the debate between idealism and realism, or between objectivism and constructivism (see for instance Clark 1997: 173). And there is also the possibility of having non-representationalism *with* realism (Zahidi 2014).

Be that as it may, before *The Embodied Mind*, Varela had proposed another argument against global representationalism, clearly pointing to the endorsement of the explicitly eliminativist trend described above. I will present and develop this argument in section III. For now, after having clarified in this section the various targets and positions of enactive anti-representationalism, I would like to present four basic *strategies* that we can find in the enactive literature and that make it possible for enactivists to defend anti-representationalist and non-representationalist positions—and especially the eliminativist trends. These strategies are often interrelated, and they can be cumulated. In the next section, I briefly present the *philosophical strategy* and the *explanatory strategy*. Section III will deal with the *ontological strategy* and with the *epistemological strategy* (the one defended by Varela (with Maturana) before *The Embodied Mind*).

II. Philosophical and explanatory strategies against representationalism

Depending on its scope of application (global, basic or local), the philosophical anti-representationalist strategy can support (10), (11), or (12), since it is a strategy against ontological representationalism. This strategy consists in arguing that "mental representation" is the wrong conceptual unit for defining the cognitive relations which take place between cognizing (or perceiving,...) organisms and their environments. This strategy is based on considerations on the (alleged) nature of cognition (if one endorses anti-representationalism for cognition in general), of perception (if one endorses anti-representationalism for perception) or of reasoning (if one endorses anti-representationalism for reasoning), etc.., and more precisely on the (alleged) nature of their intentionality. Many philosophers, from John Dewey, Martin Heidegger, Ludwig Wittgenstein, Maurice Merleau-Ponty and Emmanuel Lévinas to Hubert

Dreyfus, Charles Taylor, Hilary Putnam, Robert Brandom, John McDowell or Charles Travis have developed non-representationalist approaches of conceptual, perceptual, doxastic or epistemic intentionality. The philosophical strategy that enactivism can deploy (and has deployed) against representationalism often relies on the works of some of these authors (see for instance Varela, Thompson and Rosch 1991). Nevertheless, this strategy is not without certain defects. The main potential problem seems to be the following: this strategy is based on philosophical considerations about intentionality, cognition, knowledge or reasoning as *personal-level* phenomena. These considerations do not seem to entail a rejection of representationalism as a definition of what makes these phenomena causally possible on a subpersonal level[26]. On the contrary: representationalism, as a hypothesis concerning the existence of subpersonal (and possibly intracranial) mental representations endowed with natural content, may always be invoked for explaining for instance *how* personal-level perception, cognition or intentionality is for action or develops from practical coping or embodiment. This philosophical strategy may also be closer to *really radical enactivism* than to *radical enactivism*: the works of the philosophers mentioned above may include criticisms of the idea that non-basic forms of cognition involve *representational* content, even if the latter one is defined from social, linguistic or cultural resources.

The explanatory strategy against representationalism consists in arguing that it is not necessary to posit the existence of mental representations and to use the concept "mental representation" when one wants to explain the operations that make cognition possible. We have other explanatory resources which do not involve or presuppose an appeal to representational properties (Calvo Garzon 2008). We here find the grounds of the 7th, 8th and 9th versions of non-representationalism defined above, since the explanatory strategy is basically against methodological representationalism. Those grounds can therefore also partially justify global and basic anti-representationalisms: indeed, they will justify its methodological components. This strategy will also only partially justify global non-representationalism: the latter position concerns *all* possible cases of cognition, a level of universality that cannot be reached by empirical cases and applications of the explanatory strategy.

Local versions of this explanatory strategy may always be compatible with basic and local representationalisms (ontological and methodological, or methodological only). In order to demonstrate its viability and its global, basic or local scope, this explanatory strategy must provide a sufficient range of examples and applications. Its general (or basic) scope will often be achieved by accumulating local explanatory or predictive successes. Classical cases, today, include the engineering of artificial creatures that work without repre-

[26] This remark was already made by Rorty (1979: 230-256).

sentations at all (and not only explicit representations, as in Brooks' famous case), but also the study of insect cognition (Webb, 1994), developmental processes (Thelen & Smith 1994), motor coordination, or perceptual supplementation (Lenay & Steiner 2010). If there should be an empirically robust basic methodological non-representationalism, it will be built out of *many* cases of local methodological representationalism[27]. But this hypothetical inductive generalization is not the only challenge that is encountered by this explanatory strategy against representationalism. Indeed, according to its representationalist opponents, it must also show that

> (1) the explanatory posits it appeals to for explaining cognitive phenomena do not really have or involve representational or semantic properties. This requirement applies, for instance, to attractors in a dynamical system (Van Gelder 1995; Freeman & Skarda 1990), to coupling relations (Chemero 2009), to intentional arc and maximal grip (Dreyfus, 2002), to internal control parameters (Keijzer 2001), or to informationally sensitive responsiveness (Hutto & Myin 2013)

and that

> (2) these non-representational entities play a role in the explanation of cognitive behaviours, and not only reflex-like or stimulus-determined behaviours (or reflex-like and stimulus-determined parts of cognitive behaviours).

Still, when the proponent of representationalism expresses these two challenges, he must at least make clear what are, for him, the conditions in virtue of which a phenomenon is a representational phenomenon and a cognitive phenomenon, and in virtue of which a statement or description constitutes an explanation of a cognitive behaviour. Arguing, on *a priori* grounds, that cognition *is* necessarily representational or that any explanation of cognitive phenomena must involve the appeal to mental representations (as made out of natural content) would of course entail the victory of the proponent of representationalism even before the debate has taken place. Let us also recall that it is very easy to turn any non-representational entity or process into a representational entity or process. The hard job, for the representationalist, is to explain why the representationalist vocabulary is necessary and superior to the non-representationalist vocabulary for providing cognitive explanations.

[27] Needless to say, representationalists consider on *a priori* grounds that this challenge is a lost cause for the anti-representationalists. For instance, Wilson and Foglia (2011) write: "formulating an empirically adequate theory of intelligent behavior without appealing to representations at all (…) faces insuperable difficulties, and the idea that it is a relatively trivial matter to scale up from existing dynamic models to explain all of cognition remains wishful thinking and subject to just the problems that motivated the shift from behaviorism to cognitive science in the first place".

The explanatory strategy can also take the form of an original criticism of representationalism made by William Ramsey (2007): for Ramsey, methodological non-representationalism is already implicitly dominant in cognitive science, in spite of the representationalist propaganda. Indeed, if one pays attention to the properties in virtue of which the entities named "mental representations" play an explanatory role in many models of cognitive behaviour, one will see how much these properties are very rarely representational properties (even if the observer can ascribe them semantic properties): these properties are more elemental, since they often only take the form of indication, covariation, correlation or standing-in. If Ramsey's arguments are correct, many proclaimed "representationalist" explanations of cognitive phenomena are actually non-representational, and form a set of examples that can be exploited by the proponent of explanatory anti-representationalism[28]. Following Ramsey or Gallagher (2008), one can for instance wonder how much the so-called "minimal representations" that are defended by proponents of extended cognition like Mike Wheeler or Mark Rowlands are robust enough for having representational properties.

As said above, this explanatory strategy may be invoked for defending (10), but does not entail (10) by itself: one can be a methodological non-representationalist without endorsing ontological anti-representationalism (and thus (10)). More fundamentally, (10) is first of all a general position, while the cases put forward by the explanatory strategy are first of all local, and concern explanatory practices, not ontology.

I will now focus on two other strategies against representationalism that have been developed or are developed by enactivism: the *ontological strategy* and the *epistemological strategy*. Unlike the philosophical strategy, the ontological strategy is about mental representations themselves, and not primarily about cognition, perception or knowledge (and their intentionality). Unlike the explanatory strategy, the epistemological strategy entails ontological anti-representationalism.

[28] The only entities whose causal role is representational, according to Ramsey, are structural or simulational representations.

III. Ontological and epistemological strategies against representationalism

These two strategies support the idea that representational content[29] does not exist at a natural and subpersonal level, so that—by implication—(naturally) *contentful* physical structures do not exist. These two strategies are sufficient for excluding ontological representationalism, but are not sufficient for ruling out methodological versions of representationalism: only methodological non-representationalism (with the help of the explanatory strategy) can do that. In the enactive literature, these two strategies are included in global versions of anti-representationalism, ontological and methodological. But since these two strategies (only) concern the existence of mental representations, they could also be embedded in positions such as ontological anti-representationalism only, or ontological anti-representationalism *with* methodological representationalism (see section II).

The epistemological argument leads to the ontological elimination of mental representations (as made out of natural content) by being based on a consideration of what happens when an observer ascribes representational properties to an observed brain (be it isolated or embedded in an organism/environment system). It does not exclude the natural existence of representational content from an explanatory point of view (like the explanatory strategy), but from considerations on the pragmatics of representationalist explanations and descriptions. The conclusion is that it is *illusory to put natural content in the head* (the head being a very basic place where natural content and mental representations may be located, according to the representationalist orthodoxy). The ontological strategy reaches the conclusion that it is *hopeless to try to find natural content in the head and, more broadly, in the world*, from considerations on the problematic place of representational content in a naturalistic framework. Besides their ontological conclusions (global anti-representationalism), these two strategies share the idea that representational content can only be a product of linguistic and social practices.

Let us begin by the epistemological strategy.

The epistemological strategy, as it was notably and clearly developed by Varela and Maturana in their 1987 book *The Tree of Knowledge*, is derived from the considerations of the authors on the autonomy of living systems and on the organizational closure of the central nervous system

[29] In what follows, I will use “content”, “representational content” and “semantic content” interchangeably. (Representational) content is the content of mental representations. In the representationalist literature (including the one that is discussed and criticized by anti-representationalism), this content has semantic properties: it can be true or false, correct or incorrect; it carries meaning (and not only information).

(1987: 253). I ill not repeat these considerations here, and will instead focus on their implications for putting forward an epistemological strategy against representationalism.

Pragmatically, what is happening when one observer (a scientist) is using representationalism? We have a human person, the observer, facing an organ (the brain), or representations or models of the brain (images, data…). The observer entertains relations with the environment: he sees, feels, touches or converses *about* objects or states of affairs in the environment. When he converses about objects with his colleagues, he is producing and acting from linguistic contents, said or written in utterances. These linguistic contents notably exist in virtue of linguistic rules and conventions. The observer may believe that the brain he is observing entertains the same type of relation to the environment as the relations that he has, so that the brain would have cognitive relations with the objects the observer interacts with, by using or producing contents. But this is an illusion: unless one is under the grip of the prejudice that he/she is what his brain is or is doing, there is no reason to think that the operations of the brain—a subpersonal organ – and its relations to the environment are *like* our semantic operations and our cognitive relations with that environment. True, the brain plays a crucial causal role in the production of our cognitive and semantic behaviours, but that does not entail that it *harbours* all the dimensions of this cognitive and semantic behaviour, including its objects (as represented) and the linguistic contents from which one may think about something. Maturana and Varela indeed write:

> We as observers have access both to the nervous system and to the structure of its environment. We can thus describe the behavior of the organism as though it arose from the operation of its nervous system with representations of the environment or as an expression of some goal-oriented process. These descriptions, however, do not reflect the operation of the nervous system itself. They are good only for the purpose of communication among ourselves as observers. They are inadequate for a scientific explanation. (Maturana & Varela 1987: 132-133)

Terms such as “representation”, “memory”, “code”, or “information” occur in the space of human design and understanding. Their use for describing cerebral goings-on dramatically *abbreviates* and *over-interprets* dynamical patterns and regularities of biochemical events (Varela 1989: 7-16). Why “dramatically”? Because the observer mistakes the perspective of the brain for his very own perspective: he puts in the brain contents that only exist at some linguistic and (inter)personal level, *and turns these contents into natural entities*. William James already described this drama as follows:

> The *great* snare of the psychologist is the *confusion of his own standpoint with that of the mental fact* about which he is making his report. I shall hereafter call this the 'psychologist's fallacy' *par excellence*. For some of the mischief, here too, language is to blame. The psychologist, as we remarked above, stands outside of the mental state he speaks of. Both itself and its object are objects for him. Now

> when it is a *cognitive* state (percept, thought, concept, etc.), he ordinarily has no other way of naming it than as the thought, percept, etc., *of that object*. He himself, meanwhile, knowing the self-same object in *his* way, gets easily led to suppose that the thought, which is *of* it, knows it in the same way in which he knows it, although this is often very far from being the case. (James 1890, vol. 1, ch. VII: 196; author's italics).

Representationalism consists in the mistake of thinking that the representationalist idiom (and especially the notion of "content") is and must be something more than an idiom that can be trivially used for very metaphorically (and grossly) making sense of the brain: it would be the only or the best way of doing so, because the brain really harbours the units of this idiom; and these intracranial units would naturally represent the environment as we know or interact with it. For Maturana and Varela, the brain is not a solipsistic engine: it is an interacting part of the organism, and plays a role in the structural coupling of the organism with the environment. But it is not a *representational* engine: it does not relate to the environment as we do, or as machines that we have conceived do, in virtue of information and instructions (1987: 169):

> To an observer, the organism appears as moving proportionately in a changing environment; and he speaks of learning. To him, the structural changes that occur in the nervous system seem to correspond to the circumstances of the interactions of the organism. In terms of the nervous system's operations, however, there is only an ongoing structural drift that follows the course in which, at each instant, the structural coupling (adaptation) of the organism to its medium of interaction is conserved. (1987: 170-171).

Maturana already expressed this epistemological strategy in 1978:

> Representation, meaning, and description are notions that apply only and exclusively to the operation of living systems in a consensual domain, and are defined by an observer to refer to second-order consensual behavior. For this reason, these notions have no explanatory value for the characterization of the actual operation of living systems as autopoietic systems, even though they arise through structural coupling. (Maturana, 1978: 50; see also Maturana 1972: 23)

This argument is close to, but not identical with, a basic anti-representationalist argument in the Wittgensteinian tradition (see for instance Kenny 1989, chap. 10; Glock 2008; Descombes 2010; Bennett and Hacker 2003). According to this argument, it is a category mistake to think that natural and subpersonal phenomena of covariation or causal dependency could harbour or amount to representational or semantic properties. The latter properties are necessarily derived from linguistic and social practices, in which shareable structures (pictures, sentences, models,...) can acquire a representational status in virtue of what rule-following agents do and must do with them. The existence of mental representations as made out of natural content is here *a priori* excluded from the stipulation of necessary conditions for the existence of semantic content. These conditions will never be satisfied by a brain (even if having a brain is a necessary causal condition for being able to participate to linguistic and social practices from which representational properties emerge). Both sides (Maturana-Varela *and* the Wittgensteinian side) agree that semantic content is a product of interpersonal and normative practices. A (inter-)personal-level entity cannot be used as the general blueprint for explanations or descriptions at the subpersonal level. The Wittgensteinian argument focuses on what brains *cannot do* so that it is fallacious to ascribe representational properties to them, while Varela and Maturana focus more of what the observer *is doing* when he is using the representationalist talk for describing the brain, in order to show how representationalism rests on an illusion.

Let me now pass to the ontological strategy, as it has been recently and clearly exposed by Hutto and Myin in the chapters 4, 5 and 6 of their book *Radicalizing Enactivism. Basic Minds without Content* (in the same book, the authors also adopt an explanatory strategy against representationalism, see for instance their chapter 3)[30]. In these chapters, the authors put forward an ontological argument against representationalism, based on the current failures of the various philosophical attempts (Dretske, Millikan, Fodor,…) to naturalize representational content. It is well known that ontological representationalism has met and still meets many problems: the symbol-grounding problem, the problem of the causal efficacy of semantic properties, but also the problem of providing a naturalistic account of the content of mental representations. Concerning the last problem, Hutto and Myin remind us that neither informational theories nor teleosemantics are able to provide a satisfactory non-intentional explanation of the emergence of intentional and semantic properties (truth-conditionality, reference, intensionality) (the same could be said for resemblance-based accounts): either they beg the question by already coming with intentional notions, or they merely deliver covariation and indication, which are not sufficient for giving semantic or representational content. These failures to naturalize content entail that representationalism has no foundations in the naturalistic ontology proponents of representationalism generally assume. Unable to be integrated in the naturalistic ontology it claims to be a part of, the representationalist program would be "plagued with toxic debt, financed by loans it cannot pay back" (2013: 160). Since representational content has no place in a naturalistic ontology, there are good reasons to think it does not exist as an entity conveyed or produced by natural processes, including subpersonal and intracranial ones. For the author, the contents of our thoughts, imaginings or reasonings are not natural or subpersonal contents: they derive from the integration of our cognitive activities in socio-

[30] See my review (Steiner 2013) of the book.

cultural practices (the argument of the authors ultimately converge with the Wittgensteinian argument mentioned above).

IV. Do mental representations essentially have contents? Some limits of the ontological and epistemological strategies

It is now time to mention or propose some shortcomings of these two latter anti-representationalist strategies. I have said above that these two strategies were not sufficient for ruling out methodological representationalism: one can agree that naturally-made representational content does not exist or is an illusion and yet believe that our best explanatory policies should make use of content ascription to the brain (even if we know that nothing makes these ascriptions *true*). This is not an objection to these strategies, since they are explicitly (and only) against ontological representationalism.

These two critical strategies include (but are not restricted to) a common inference, ending with an eliminativist conclusion:

> P1. In the representationalist ontology, the subpersonal and intracranial phenomena named "mental representations" naturally (or intrinsically) have contents[31]
>
> P2. There is no natural (or intrinsic) content at the level of subpersonal and intracranial phenomena
>
> C. Subpersonal and intracranial mental representations, as they are conceived by the representationalist ontology, do not exist

Of course, they will justify P2 differently; this is why they are different strategies: for the epistemological strategy, there is no natural representational content because it is an illusion or an artifact; for the ontological strategy, there is no natural representational content because there is no satisfying naturalist account of representational content. Before criticizing this inference and thus these two strategies together, let me first express an objection against the justification of P2 that is proposed by the ontological strategy. Hutto and Myin's ontological strategy starts from a current state of affairs (the failures of attempts to naturalize content), and infer some general truth of it. But the fact that content has not been naturaliz*ed* until now does not mean it is not naturaliz*able*: past failures do not necessarily entail future failures.

[31] As said above, for vehicle-internalist versions of representationalism, mental representations necessarily consist in intracranial and subpersonal processed endowed with natural content. For vehicle-externalist versions, mental representations are not necessarily intracranial or endowed with natural content, but *some of them are intracranial and endowed with natural content*: this is the main claim of representationalism as I have defined it in the introduction, and as it figures in the premise 1 above.

Let me now criticize the general inference as it is endorsed by the ontological strategy and by the epistemological strategy. It is more precisely the uncritical endorsement of P1 that can be questioned. P1 can be rephrased as follows: "it is part of the concept 'mental representation' as it is used by representationalists for denoting subpersonal and intracranial processes that mental representations naturally have contents". As I have already said in the introductory section, the criticism and the demise of representationalism realized by the epistemological and by the ontological strategies is thus the criticism and the demise of a certain kind of representationalism: representationalism according to which mental representations naturally have contents. A representationalism which would come with the idea that representational content is necessarily made out of linguistic and social resources (so that there is no natural content) would not constitute a target for enactive anti-representationalism. A different but important debate would be to know to which extent basic cognition is representational… or not, when by "representational" it is now meant "involving the existence of linguistic and symbolic practices": this debate would not be a debate between representationalism and anti-representationalism as we have defined them in this paper.

If representationalists deny that mental representations have *natural* content, their representationalism is not a problem for radical enactivism. But they may also deny that mental representations have *content*, while still seeing them as natural, subpersonal and intracranial phenomena… now untouched by the ontological and the epistemological strategies. Indeed, P1 can be refused by versions of representationalism which consider that mental representations should not be modeled on cases of external representations, and thus do not entertain referential relations with their objects in virtue of some content (be it natural or not). For these versions of representationalism, mental representations, as natural, subpersonal and intracranial phenomena exist; but they do not have content. Otherwise put: the epistemological and the ontological strategies are working if one assumes that the alleged nature of mental representations is exhausted by the descriptions one can find in popular scientific accounts (but also folk accounts) of the term "representation": in these descriptions, representations have contents. But these strategies may not work if one considers that the concept "mental representation", whatever our definitions and descriptions of mental representations may be, robustly refers to a cluster of properties which are probably very different from the properties that are assigned by folk accounts of representations. Before seeing how this alternative is possible, let us pause for a moment for seeing that this objection is a classical objection that is faced by every kind of eliminativism concerning theoretical terms (Stich 1996, chap. 1). Michael Devitt summarizes the objection as follows:

> Consider how, in general, we argue directly for the *non*existence of *F*s. On the basis of the established view of *F*s, we start, implicitly if not explicitly, with an assumption about the nature of being *F*: something would not be an *F unless it were G*; being *G* is part of the very essence of being *F*. Then we argue that nothing is *G*. So, there are no *F*s. But suppose that someone responds by denying the essentialist assumption upon which the argument rests. '*F*s do not have to be *G*, they are just mistakenly thought to be *G*. So the argument proves nothing' (2009: 57; author's emphasis).

A typical eliminativist argument proceeds as follows: a theoretical term T refers to whatever entities satisfy a description (or a cluster of descriptions) generally associated with the term in a theory (the meanings of theoretical terms are defined by reference to causal roles specified by the theory). If nothing satisfies this description (i.e. if nothing instantiates the causal roles defined by the theory), there are good reasons to think that the theoretical term does not refer to anything (and not only that the theory is false): hence the *justified* fates of the terms "phlogiston", "crystalline spheres", "caloric", "aether"… and maybe "mental representation" in the history of science. This line of reasoning assumes a descriptivist theory of reference, according to which the reference of a term is determined by the descriptions associated with the term. But once one endorses a causal theory of reference, things get a little bit more complicated. A descriptivist theory of reference is very suitable for explaining why theoretical terms of the past have been abandoned, but may fail for explaining how some theoretical terms have been retained throughout history in spite of massive changes of meaning. According to a causal theory of reference, the reference of a term is not a matter of senses or descriptions, but of a causal-historical chain between the term and its referent. The descriptions associated with a term may be false, and yet the term might refer to some event or property. A term might refer to something whose key properties are not the ones mentioned in its current intension. The causal theory of reference easily explains why terms such as "planet", "atom" or "gene" have been correctly maintained (and not eliminated) in the course of history even though their meanings (and the theories they have been included in) have deeply changed. Why should it not also be the case for the concept "mental representation" as used for denoting a natural, intracranial and subpersonal phenomenon?

What are the conditions in virtue of which one can say that some entity or property does not exist, rather than say that it exists, although it is very different from what one thought and thinks about it? What are the conditions in virtue of which some term does not refer to anything, rather than referring to something which is very different from what the descriptions associated with it prescribe? There is no definite answer(s) to these questions: a consideration of the properties of the theory in which the term is defined may be crucial, but these properties can be and have been very different from case to case. The result (that is, the elimination or the conservation of the term) can derive from the fact the theory is seen as true or false, too simple or too complicated, necessary (being the only game in town) or useless, reducible or not to a new theory… but it can also be generated by social and pragmatic factors. It is very easy to assert, *ex post facto*, that eliminated scientific concepts *had to* be eliminated because right from the start, nothing corresponded to them. But at the time of the controversy between their proponents and their opponents, this alleged absence of reference was exactly the disputed issue: it was a disputed argument, not an undisputed conclusion. What firstly motivated the effective elimination of (pseudo)scientific concepts was generally not the inexistence of their referents (how could something inexistent make by itself a causal difference?), and not only the availability and the relevance of other concepts ("oxygen" over "phlogiston" for instance), for the concepts that finally became eliminated were available and theoretically relevant: it was also a set of interests, purposes and institutional factors which increasingly led to their elimination. A classical proponent of eliminativism, Patricia Churchland, was clearly aware of the heterogeneity of the causes that can contribute—or not—to the elimination of a theory or of a theoretical term when she wrote that

> The whim of the central investigators, the degree to which confusion will result from retention of the old terms, the desire to preserve or to break with past habits of thought, the related opportunities for publicizing the theory, cadging grants, and attracting disciples all enter into decisions concerning whether to claim identities and therewith retention or whether to make the more radical claim of displacement. (Churchland 1986: 283-284).

Knowing that there is no historical law or methodological rules that would allow us to directly infer the elimination of a term from the failure of the theory (or of the description) in which (or with which) it is defined, proponents of anti-representationalism should not think that the demise of natural representational content entails the end of mental representations as natural entities. The concept "mental representation" might actually refer to entities without content *and be retained* in spite of massive semantic changes. This is not a vague theoretical possibility: this change of perspective is already in place if we consider some minor (yet existing) versions of representationalism.

Taking into account (like Hutto & Myin!) the massive failure of attempts to propose a naturalization of the semantic relation that is supposed to exist between mental representations and environmental facts and properties, Dan Lloyd has proposed to consider the property of mental representation as a monadic property:

> Every attempt to express the relation of representation in non-represen-tational terms has failed to meet the constraints of content, being either too short in range or too wide in focus. If neither dyadic nor polyadic relations are adequate for the task, then it is perhaps time for a proposal of last resort: what if repre-

> sentation is not a relational property at all, but rather a monadic, nonrelational property? (Lloyd 2003: 938).

This definitional move is already sufficient for not being shaken by Hutto and Myin's criticism of representationalism, since the latter criticism assumes that the concept of "mental representation" refers to a dyadic (at least) property. Mental representations, here, are not relational properties: they do not refer to anything. The requirement of naturalizing their referential or intentional dimension is therefore dropped off.

But one can go further, and defend representationalism by getting rid of the very reality of content (at least as a reference-enabling entity). As France Egan (2010) has clearly suggested, most versions of representationalism endorse what she calls the "Essential distal content view", made of the three following commitments:

> 1. Mental representations are distally interpreted: they are about objects and properties in the environment.
>
> 2. Computational states and processes are type-individuated with reference to these distal objects and properties[32]
>
> 3. The relation between mental representations and the distal objects and properties to which they are mapped is a substantive, naturalistically specifiable relation.

Based on the analysis of Marr's theory of early vision and on a computational theory of motor control, Egan (2014) suggests that it is possible to endorse representationalism without endorsing these three commitments: distally-defined content does not need to play an individuating and a causal role in computational models of cognitive tasks. But it might still have an *explanatory* role. Its ascription to some internal states would be necessary for explaining how a computational process is the exercise of a *cognitive* capacity:

> A semantic interpretation of a computational mechanism is necessary to explain how a formally characterized process, in a certain context (say, when connected to certain performance systems, or situated in a certain external environment) constitutes the exercise of a *cognitive* capacity, such as computing the depth of the scene, or the syntactic structure of an acoustic input. (Egan, 2003: 100; author's emphasis).

[32] This second commitment shows that Egan defines these commitments as belonging to the computational tradition. I propose here to attribute these commitments to other representationalist traditions as well. This can be done by rephrasing the second commitment into "distal objects and properties are mentioned in order to type-individuate most cognitive states and processes".

Egan's proposal is both progressive and conservative: progressive, in so far as she argues that distal content does not have causal and individuating roles in computational models of cognitive processing. For this reason, her version of representationalism avoids most of the criticisms of the ontological and of the epistemological strategies: the relation between mental representations and distal objects is not a real and naturalistically specifiable relation, since distal content is ascribed to these mental representations. This ascription of distal content is based on the fact these internal states and structures co-vary with environmental properties and facts; but these states and structures do not represent their normal distal cause (Egan, 2010: 257). Still, Egan's position is conservative with respect to the notion of content: even *unreal* (that is, not grounded on a natural relation, and not having a causal and individuating role) and defined as the product of an ascription, content must be retained[33]. It is the ascription of content which defines what a *cognitive* mechanism or task is. (The ascription of) content is therefore methodologically necessary. Mental representations exist (they are not ascribed), although their content is (only) ascribed: the explanatory strategy can criticize the latter claim, while the ontological strategies seen above cannot attack the first claim.

Chomsky is much more radical than Egan, by defending the idea that content *tout court*, including the concept of "content", has no place at all in cognitive science[34]. Indeed, for a long time now, Chomsky is convinced that folk and philosophical notions like "content", "intentionality" and "reference" have no place at all in a naturalistic framework dedicated to the understanding of cognitive faculties. There is thus no need to discuss the current failures of attempts to naturalize representational content for dropping out content in the definition of mental representations:

> The central problem that troubles me is this. I do not know of any notion of 'representational content' that is clear enough to be invoked in accounts of how internal computational systems enter into the life of the organism. And to the extent that I can grasp what is intended, it seems to be very questionable that it points to a profitable path to pursue. (Chomsky 2003a: 274).
>
> To be an Intentional Realist, it would seem, is about as reasonable as being a Desk- or Sound-of-Language- or Cat- or Matter-Realist; not that there are no such things as desks, etc., but that in the domain where questions of realism arise in a serious way, in the context of the search for laws of nature, objects are not conceived from the peculiar perspectives provided by concepts of common-sense. (...) Intentional phenomena relate to people and what they do as viewed

[33] Of course, Egan's criticism of content as a distally defined entity leaves intact the possibility that narrow content exists.

[34] See also Jackendoff (1992, chap. 8) for a similar (and developed) claim and, of course, Stich's (1983, chap.8) claim that psychology (under the form of a syntactic theory of the mind) has no need to postulate content, semantic properties or truth conditions.

> from the standpoint of human interests and unreflective thought, and thus will not (so viewed) fall within naturalistic theory, which seeks to set such factors aside. Like falling bodies, or the heavens, or liquids, a "particular intentional phenomenon" may be associated with some amorphous region in a highly intricate and shifting space of human interests and concerns. But these are not appropriate concepts for naturalistic inquiry (…). If 'cognitive science' is taken to be concerned with intentional attribution, it may turn out to be an interesting pursuit (as literature is), but is not likely to provide explanatory theory or to be integrated into the natural sciences. (Chomsky 2000: 21-23).

Hutto & Myin might well agree with Chomsky when he writes that "naturalistic inquiry will always fall short of intentionality" (2000: 45); but whereas they will take this as a good argument for eliminating mental representation as made out of natural content, Chomsky will take it as a good reason for divorcing mental representations from content and intentionality. Content is not defined by functional role or reduced to formal properties, and it does not play any causal or explanatory role, *for there is no content here, even conceived as "narrow content".* The assumption that mental representations have content is, from a naturalistic point of view, a useless and eccentric assumption. Any ontological, epistemological, philosophical or explanatory query concerning content will not be a problem for Chomsky's representationalism. This variety of representationalism is immune to the accusation of resting upon an epistemological fallacy and to the accusation of resting upon a non-existent naturalistic theory of content. For Chomsky, mental representations are individuated from their role in cognitive processing. The functional roles of mental representations are here related to properties that have nothing to do with content, truth conditions, reference, or intentionality. Their important properties are formal or syntactic. These formal and syntactical properties are sufficient for individuating and studying the causal role of these representations in cognitive processing. These representations do not mean or represent anything; defining their reference is of no scientific interest:

> The internalist study of language also speaks of "representations" of various kinds, including phonetic and semantic representations at the "interface" with other systems. But here too we need not ponder what is represented, seeking some objective construction from sounds or things. (…) Accessed by performance systems, the internal representations of language enter into interpretation, thought, and action, *but there is no reason to seek any other relation to the world, as might be suggested by a well-known philosophical tradition and inappropriate analogies from informal usage*. (Chomsky 1995: 53 ; my emphasis).

"Informal usage", here, means the very widespread tendency to embrace a linguistic model of mental representations, assuming they have semantic content or truth conditions, like daily linguistic products. The proponent of contentless representationalism, here, wants to sever the ties between the explanatory posits of cognitive science and our folk understanding of what representations are, but considers that the very term of "representation" can still be applied for describing the natural, intracranial and subpersonal reality he is studying (even though its meaning, in scientific inquiry, is very different from our daily conceptions). Contentless representations entertain functional relations with external phenomena: they occur when and only when the organism interacts or deals with these external phenomena. In this sense, a "number-representation" is a representation of a different functional type than a "face-representation", but is not to be defined as a representation *of* an external item. But—and *pace* Egan—content-ascription is not a necessary methodological stance. At the very best, content-ascription can play some auxiliary role in the informal presentation of a computational theory, but not within the computational model itself (Jacob 2010: 231):

> There is no meaningful question about the "content" of the internal representations of a person seeing a cube under the conditions of the experiments, or if the retina is stimulated by a rotating cube, or by a video of a rotating cube; or about the content of a frog's "representation of" a fly or of a moving dot in the standard experimental studies of frog vision. No notion like "content," or "representation of" figures within the theory, so there are no answers to be given as to their nature. (Chomsky 1995: 52).

Chomsky's representationalism aims at doing without content (hence my use of "contentless representationalism" for naming this version of representationalism) [35]. Some commentators like Georges Rey (2003a, 2003b) have argued that his clarifications of what he consequently meant by "representation" could not escape reference to intentional properties, and thus content. For instance, there are places in which Chomsky equates "represent" with "implement" (2003: 276): of course, this definition saves "representation" from referential properties, but seems to presuppose that some intentionally characterized item—that is, content!—is implemented. This is an instructive debate I will not consider here. More broadly, as said in the introduction, I do not consider at all that contentless representationalism is a refutation of enactive anti-representationalism, so that we should embrace or develop it. It is enough for the purpose of this section to note that contentless representationalism is immune to the criticisms made by proponents of the ontological and the epistemological strategies against representationalism, since it does not rest upon the assumption that natural (and especially intracranial and subpersonal) representational content exists, or that its ascription is necessary.

A table may be helpful here in order to underline the differences between contentless representationalism, representationalism, and enactive representationalism, but also some common points between enactive anti-representationalism and classical representationalism: as said from the introductory section, enactive anti-representationalism (paradigmatically: Hutto

[35] See Collins (2007) who argues that this position has been exemplified by Chomsky since his seminal critical review of Skinner's *Verbal Behavior* in 1959.

and Myin's radical enactivism, and also Maturana and Varela's criticism of representationalism) is *against* mental representations made out of natural content, but not against mental representations having non-natural content: content is a product of socio-cultural practices. This criticism of natural content is enough for understanding how enactive anti-representationalism is radical, but we must not overlook the possibility there are proximities between enactive anti-representationalism and classical representationalism on other issues: both sides presuppose that "content" and "representation" stand or fall together (if some entity does not have content (be it natural or not), it is not a representational entity), and (less importantly) both sides consider that content (be it natural or not) can play an explanatory role in cognitive science.

	Is there natural content (so that there might be mental representations as physical structures carrying natural content)?	Is "content" essential to the notion of "representation"?	Should content (natural or not) play an explanatory role in cognitive science?
Classical (content) representationalism	Yes	Yes	Yes
Enactive anti-representationalism	No	Yes	Yes (for non-basic cognition)
Chomsky's contentless representationalism	No	No	No

Table 3

V. Concluding Remarks

In the spirit of this paper, I take contentless representationalism as an opportunity, not to reject, complement or even amend enactive anti-representationalism, but to reflect on some presuppositions and challenges of enactive anti-representationalism. Contentless representationalism invites us to retain representationalism (or at least the claim that there are subpersonal and intracranial phenomena that are naturally representational) if and only if we accept that mental representations have no content, no truth and satisfaction conditions, no reference, no intentionality, and no definitional role for cognition. The concept "mental representation" is conserved, but its meaning is deeply changed. In this sense, contentless representationalism is an alternative to anti-representationalism, which invites us to abandon the concept "mental representation" for describing or modeling natural phenomena such as neural processes, since there is no natural content at the level of intracranial and subpersonal structures. Anti-representationalism encourages the use of concepts such as indication, covariation, or correlation for defining the activity of neural patterns. Still, for contentless representationalism, it might turn out that the concept "mental representation" refers to some kind of inner states which are currently described by anti-representationalists with these latter concepts. These states have neither content nor intentionality, but play a role in enabling the achievement of a cognitive task whose cognitive character may be defined by explanatory conventions, or by taking into account the integration of the achievement of the task in broader organism-environment interactions (if one endorses vehicle-externalism). The proponent of anti-representationalism will probably object to contentless representationalism: "WHY still go on calling these entities 'representations' if they do not have content, reference or intentionality?" Is not "contentless representation" a pure oxymoron, as Hutto & Myin suggest (2013: 84)? Two replies might be proposed—I see both of them as pointing towards challenging issues for enactive anti-representationalism:

1) This question will also be asked by the proponent of content representationalism. This fact is revealing: as shown above, enactive anti-representationalists *and* classical representationalists may share a common assumption, the idea that any kind of representation *must* have content (be it natural or non-natural). Both sides agree that (representational) content and (representational) vehicle stand or fall together. Quite ironically, (failed) attempts to naturalize content *and* the ontological and epistemological strategies against representationalism could derive from a common confusion: defining or eliminating a scientific concept—"mental representation"—from the resources of common sense, where basic cases of representation do have content. If one makes that confusion, it becomes natural to think that the fate of the concept "mental representation" is linked to the fate of "mental content".

Anti-representationalist enactivists do not seem to see that there may be scientific changes by which we *retain* concepts even though the meanings of the latter ones are changing. Or, at least, and in better words: they do not accept this possibility for the concept of "mental representation" as denoting natural, intracranial and subpersonal phenomena. But, in this case, it would be interesting to know why this possibility is refused to "mental representation" while it has been accepted for "genes" (Fox-Keller, 2002) or "atoms" (Pullman, 1998): their meanings (and the theories they figure in) have deeply changed, and yet the concepts (or minimally, the terms) have been retained[36]. As Paul Griffiths

[36] Here is what E.A. Carlon already remarked in 1966: "The gene has been considered to be an undefined unit, a unit-character, unit-factor, a factor, an abstract point on a recombination map, a three-dimensional segment of an anaphase chromosome, a linear segment of an interphase chromosome, a sac of genome's, a series of near sub-genes, a spherical unit defined by a target theory, a dynamic functional quantity of one specific unit, a pseudoallele, a specific chromosome segment subject to position effect, a rearrangement within a continuous chromosome molecule, a cistron within which fine structure be demonstrated, and a linear segment of nucleic acid specify-

and Karola Stotz (2007) clearly showed, the changes of meaning of the concept "gene" emerged from a dialectics between a structural conception (anchored in biochemistry) and a functional conception (based on the observable results of hybridization between DNA molecules) of the gene. The search for the lower-level mechanisms fulfilling the functional role attributed to the gene led to gradual changes in the definition of this functional role: contemporary genes are not the fundamental units of mutation, of replication or of recombination; and they cannot be identified with DNA-segments that would unequivocally "code for" proteins or "determine" phenotypic traits. Various eliminativist temptations were fostered during this process of refinement of the functional role of the gene with regard to the discovery of its material realization, but the concept "gene" was retained (as said above, only naive presentism and realism would lead us to explain that conservation of the concept "gene" by mentioning the "real existence of genes"). Contemporary "atoms" are not indestructible, indivisible and immutable, and atoms of the same element do not necessarily have identical properties and mass.

Arguing that the concept of "mental representation" should be eliminated because—unlike "genes" and "atoms"—it does not refer to anything would be begging the question, because this assumption of non-existence is based on the idea that natural content does not exist…which is precisely the clause that is modified in the new intension of the concept "mental representation"! A possibly better answer to the question "why drop the concept 'mental representation' just because of massive change of meaning?" might be that the concept of "mental representation" *has to* be connected to our daily, pre-scientific concept of "representation" (which essentially includes "content" in its intension), so that giving up this connection is renouncing to some intelligible concept of "mental representation". But there is no *a priori reason* to think that concepts proposed in the context of scientific inquiry have to be derived or must respect the bounds of common sense (where representations have content). On the contrary: one may think that the autonomy of these concepts is the best guarantee for avoiding the fallacies correctly diagnosed and criticized by Varela & Maturana,… and by Chomsky: using concepts such as "content" with their common sense meaning for describing and explaining intracranial and subpersonal processes, as so many forms of representationalism do. But once this diagnosis is made, nothing forces scientists (like Chomsky) not to use concepts by changing their meanings, cutting all ties with common sense (for instance: mental representations with no content)[37]. Only ordinary-language

ing structural or regulatory product" (Carlson E.A., *The Gene: A Critical History*, Philadelphia/London, Saunders, 1966: 259 ; quoted by Burian 1985: 5))

[37] Godfrey-Smith (2004: 159) suggests: "When a cognitive scientist works on mental representation, what we often find is a special kind of meeting two conceptual frameworks and mindsets (…) We should think of the 'representational' concepts used in cognitive science as *amalgams*, or *hybrids*, born of the interaction between the ordinary interpretative habits that cognitive scien-

foundationalism can lead us to think that if scientists use concepts that are formally identical with the concepts of common-sense (like “space”, “life”, “mass”, “negative charge”, “matter”, “field” and “representation”) then these scientific concepts *must* have and keep the same meaning as the concepts of common sense, even if this raises the risk of committing category mistakes (at the benefit of armchair critiques of scientific theories).

2) Instead of “contentless representations’, anti-representationalist enactivism will encourage the use of other concepts, such as “indication”, “informationally sensitive responsiveness” “covariation”, or “correlation” for describing what occurs at the natural, subpersonal and intracranial level of cognitive processing. Semantically and logically, there are all the differences in the world between these concepts and “representation”, if “representation” comes with the very idea of content. Still, pragmatically—if one considers the contexts and the functions of their uses –, all these concepts will be analogous to representational concepts such as “mental representation” or “information-carrying structures” *as long as anti-representationalists do not state how the use of these concepts should be embedded in new ways of building and assessing cognitive explananda and explanantia, especially those pertaining to neural processes.* Otherwise put: it would be very naive to suppose that changing a word for another is equivalent to a conceptual change: concepts, their meanings and their uses are always embedded in theoretical enterprises, interests and programs. Changing the word “represent” for another word (“indicate”, “is correlated with”, “is informationally sensitive to”,…) without changing the theoretical context is *not* a conceptual change; it is just a linguistic trick. The proponent of anti-representationalism may legitimately object to contentless representationalism that one does not see the point of retaining the concept of “representation” once it has undergone so many semantic changes; but the proponent of contentless representationalism may also legitimately reply that the anti-representationalist is guilty of not providing a sufficiently clear alternative to the use of the concept “representation”, beyond a mere change of words. As long as we will not be able to imagine how we can *do* cognitive (and especially neural) science in a non-representationalist framework, it is unlikely non-representational words will be able to perform a real non-representational job. We know that ‘representation’ is not a neutral word: it

tists have just in virtue of being people, and the scientific aims of describing precise, naturalistic and empirically studiable relations between organisms and environments. The representational concepts used in cognitive science are products of marriages between folk semantics concepts and a family of naturalistic concepts of physical specificity-concepts of connection and directedness that are based on causal, nomic and functional concepts.” (author’s emphasis). What I am writing here suggests that the marriage between folk semantics and scientific concepts, in the case of representational concepts in cognitive science, has not been and is not an equal marriage: folk concepts still wear the trousers. True, mental representations can be subpersonal and nonconceptual; but they are seen as being necessarily contentful.

comes with a specific way of thinking, describing and explaining cognition, and more precisely the role of the substrate of these “mental representations”—namely, the brain—in cognitive processes. Symmetrically, if a new word should replace “mental representation”, it must come with new ways of thinking, describing and explaining cognition and brain activity. *Of course*, the enactive proponent of anti-representationalism can already provide some sketches of radically new theoretical contexts for the use of a non-representationalist vocabulary: works and developments on autopoiesis, autonomy, coupling or sense-making can help framing this context. Reconsidering what external representations are, and what cognitive jobs they can do so that we may have contentful thoughts in an anti-representationalist framework, is also an urgent task[38], along with a reconsideration of how the “representational hunger” of “off-line” and “abstract” cognitive tasks may be satisfied (Degenaar and Myin, forthcoming). But, as said at the very beginning of this text, enactive anti-representationalism and attempts to construct positively an enactive cognitive science are nowadays quite separated. We should only hope that this separation is very temporary and contingent. But it is also possible that this separation is the symptom of the existence of different irreducible commitments—in the enactive framework—to what is centrally required for the overcoming of classical cognitive science. Some might want to develop and to refine an explicit eliminativist stance on mental representations, while others might think that the age of representation is so over that one should not lose one’s time arguing with the past. Even if the first strategy faces some theoretical difficulties I have outlined in this paper, I believe it would be a mistake to abandon it: I do not see how the second strategy would be *a priori* immune against all forms of representationalism. And besides being parts of the same research program (namely, enactivism), these two strategies can converge on several crucial issues such as the status of meaning and content in nature (including our second nature), or the theoretical place that must be allocated to the brain in the definition and explanation of cognitive processing.

References

Bennett, M. R. and Hacker, P. M. S. 2003. *Philosophical Foundations of Neuroscience*. Wiley-Blackwell.

Bickhard, M. H. 2009. Interactivism: a manifesto. *New Ideas in Psychology,* 27: 85-95.

Burian, R. M. 1985. On conceptual change: the case of the gene. D. J. Depew & B. H. Weber, *Evolution at a Crossroads – The New Biology and the New Philosophy of Science* (21-42). Cambridge (Mass.): MIT Press.

[38] I have proposed some suggestions in Steiner (2014b).

Calvo Garzon, F. 2008. Towards a general theory of antirepresentationalism. *Brit. J. Phil. Sci.*, 59: 259-292.

Chemero, A. 2000. Anti-representationalism and the dynamical stance. *Philosophy of Science*, 67: 625-647.

Chemero, A. 2009. *Radical embodied cognitive science*. Cambridge (MA)/London: MIT Press.

Chomsky, N. 1995. Language and nature. *Mind* 104 (413): 1-61.

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*, Cambridge University Press.

Chomsky, N. 2003a. Reply to Egan. L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (268-274). Blackwell.

Chomsky, N. 2003b. Reply to Rey. L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (274-287). Blackwell.

Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge (MA): MIT Press.

Clancey, W. J. 1997. *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge University Press.

Clark, A. 1997. *Being There. Putting Brain, Body and World Together Again*, Cambridge (MA): MIT Pess.

Clark, A. 2001. *Mindware*. New York and Oxford: Oxford University Press.

Clark, A. 2008. *Supersizing the Mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.

Collins, J. 2007. Meta-scientific eliminativism: a reconsideration of Chomsky's review of Skinner's *Verbal Behavior*. *Brit. J. Phil. Sci.*, 58: 625–658.

Degenaar, J. and Myin, E. Representation-hunger reconsidered. Forthcoming in *Synthese*.

Descombes, V. 2010. *The Mind's Provisions: A Critique of Cognitivism*. S. Schwartz transl. Princeton University Press.

Devitt, M. 2009. On determining what there isn't. D. Murphy and M. Bishop, eds. *Stich and His Critics* (46-61). Oxford-Malden, Blackwell.

Dreyfus, H. 2002. Intelligence without representation – Merleau-Ponty's critique of mental representation. The relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences*, 1: 367–383.

Egan, F. 2003. Naturalistic inquiry: where does mental representation fit in? L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (89-104). Blackwell.

Egan, F. 2014 How to think about mental content. *Philosophical Studies*, 170: 115-135.

Egan, F. 2010. Computational models: a modest role for content. *Studies in History and Philosophy of Science* (special issue on Computation and Cognitive Science), 41: 253-259.

Ellis, R. and Newton, N. 2010. *How the Mind uses the Brain*. Open Court.

Fodor, J. 1987. *Psychosemantics. The Problem of Meaning in the Philosophy of Mind.* Cambridge (MA): MIT Press.

Foglia, L. and Grush, R. 2011. The limitations of a purely enactive (non-representational) account of imagery. *Journal of Consciousness Studies,* 18 (5-6): 35-43.

Fox Keller, E. 2002. *The Century of the Gene*. Cambridge (MA): Harvard University Press.

Freeman, W. & Skarda, C. 1990. Representations: who needs them? J. McGaugh, N. Weinberger, and G. Lynch, eds. *Brain Organization and Memory: Cells, Systems and Circuits* (375-380). Oxford: Oxford University Press.

Gallagher, S. 2008. Are minimal representations still representations? *International Journal of Philosophical Studies,* 16 (3): 351-69.

Glock, H. -J. 2008. Concepts: representations or abilities? E. Di Nucci and C. McHugh, eds. *Content, Consciousness and Perception: Essays in Contemporary Philosophy of Mind* (36-61). Cambridge Scholars Publishing.

Godfrey-Smith, P. 2004. On folk psychology and mental representation. H. Clapin, Ph. Staines and P. Slezak, eds. *Representation in Mind: New Approaches to Mental Representation* (147-162). Amsterdam: Elsevier

Griffiths, P. and Stotz, K. 2007. Gene. D. L. Hull and Michael Ruse, eds. *The Cambridge Companion to the Philosophy of Biology* (85-102). Cambridge/New York: Cambridge University Press.

Haselager, P., de Groot, A. and van Rappard, H. 2003. Representationalism vs. anti-representationalism: a debate for the sake of appearance. *Philosophical Psychology*, 16 (1): 5-23.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge (MA): MIT Press.

Hutto, D. & Myin, E. 2013. *Radicalizing Enactivism. Basic Minds Without Content*. Cambridge (MA)/London: MIT Press.

Jackendoff, R. 1992. *Languages of the Mind*. Cambridge (MA): MIT Press.

Jacob, P 2010. . The scope and limits of Chomsky's naturalism. J. Bricmont and J. Franck, eds. *Chomsky Notebook* (211-234). New York: Columbia University Press.

James, W. 1890. *The Principles of Psychology*. New York: Dover.

Keijzer, F. 2001. *Representation and Behavior*. Cambridge (MA): MIT Press.

Kenny, A. 1989. *The Metaphysics of Mind*. Oxford-New York: Oxford University Press.

Lenay, Ch. and Steiner, P. 2010. Beyond the internalism/externalism debate: the constitution of the space of perception. *Consciousness and Cognition*, 19: 938-952.

Lloyd, D. 2003. Representation, Philosophical issues about. Encyclopedia of Cognitive Science. NY, McMillan.

Maturana, H. 1972. Biology of cognition. H. Maturana and F. Varela, *Autopoiesis and Cognition. The Realization of the Living* (1-58). Dordrecht – Boston: Reidel.

Maturana,H. 1978. Biology of language: the epistemology of reality. G. Miller and E. Lenneberg, eds. *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg* (27-63). New York: Academic Press.

Maturana, H. and Varela, F. 1987. *The Tree of Knowledge. The Biological Roots of Human Understanding*. Boston and London: Shambhala.

Menary, R. 2007. *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave-Macmillan.

Noë, A. 2004. *Action in Perception*. Cambridge (MA): MIT Press.

O'Regan, J. K. and Noë, A. 2001, A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences,* 24: 939-1031.

Popper, K. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Pullman, B. 1998. *The Atom in the History of Human Thought*. A. Reisinger (transl.). Oxford University Press.

Ramsey, W. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press.

Rey, G. 2003. Chomsky, intentionality, and a CRTT. L. Antony & N. Hornstein, eds. *Chomsky and His Critics* (105-139). Oxford: Blackwell.

Rey, G. 2003b. Representational content and a Chomskyan linguistics. A. Barber, ed. *Epistemology of Language* (140-186). Oxford University Press.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton University Press.

Rowlands, M. 2010. *The New Science of the Mind. From Extended Mind to Embodied Phenomenology*. Cambridge (MA)/London: MIT Press.

Sellars, W. 1956. Empiricism and the philosophy of mind. H. Feigl and M. Schriven, eds. *Minnesota Studies in the Philosophy of Science, Volume I: The Foundations of Science and the Concepts of Psychology and Psychoanalysis* (253-329). University of Minnesota Press.

Shapiro, 2014. Review of *Radicalizing Enactivism. Basic Minds without Content*, *Mind*, forthcoming.

Sprevak, M. 2013. Fictionalism about neural representations. *The Monist*, 96 (4): 539-560.

Steiner, P. 2010. The bounds of representation. A non-representationalist use of the resources of the model of extended cognition. *Pragmatics & Cognition,* 18 (2): 235-272.

Steiner, P. 2011. Enaction, pragmatisme et minimalisme représentationnel. N. Bault, V. Chambon, N. Maïonchi-Pino, F.-X. Pénicaud, B. Putois and J.-M. Roy, eds. *Peut-on se passer de representations en sciences cognitives?* (191-210). Bruxelles-Paris: De Boeck

Steiner, P. 2013. Les radicaux de la cognition. A propos de *Radicalizing Enactivism. Basic Minds without Content (2013)*. *Intellectica,* 59, 2013/1: 299-308.

Steiner, P. 2014. A problem for representationalist versions of extended cognition. *Philosophical Psychology*, forthcoming.

Steiner, P. 2014b. The delocalized mind. Judgements, vehicles, and persons. *Phenomenology and the Cognitive Sciences,* 13 (3): 437-460

Stewart, J. 1996. Cognition = Life: implications for higher-level cognition. *Behavioural Processes,* 35: 311-326.

Stewart, J., Gapenne, O. and Di Paolo, E. A., eds. 2010. *Enaction. Towards a New Paradigm for Cognitive Science*. Cambridge (MA): MIT Press.

Stich, S. 1983. *From Folk Psychology to Cognitive Science: The Case against Belief*. Cambridge (MA): MIT Press

Stich, S. 1996. *Deconstructing the Mind*. Oxford-New York: Oxford University Press.

Sutton, J. 2010. Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. R. Menary, ed. *The extended mind* (189-225). Cambridge (MA)/London: MIT Press.

Thelen, E. and Smith, L. B. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge (MA): MIT Press.

Thompson , E. 2011. Replies to commentaries. *Journal of Consciousness Studies* 18 (5/6): 76-223.

Van Gelder, T. 1995. What might cognition be if not computation ? *Journal of Philosophy*, 92: 345-381.

Varela, F. 1989. *Autonomie et connaissance. Essai sur le vivant*. Paris: Seuil.

Varela, F., Thompson, E. & Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge (MA): MIT Press.

Webb, B. 1994. Robotic experiments in cricket phonotaxis. D. Cliff, P. Husbands, J.-A. Meyer and S. W. Wilson, eds. *From animal to animats 3. Proceedings of the third international conference on simulation of adaptive behaviour*. Cambridge (MA): MIT Press.

Wheeler, M. 2005. *Reconstructing the Cognitive World. The Next Step*. Cambridge (MA): MIT Press.

Wilson, R. 2004. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge: Cambridge University Press.

Wilson, R. and Foglia, L. 2011. Embodied Cognition. *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition). Edward N. Zalta, ed. URL = <http://plato. stanford. edu/archives/fall2011/entries/embodied-cognition/>.

Zahidi, K. 2014. Non-representationalist cognitive science and realism. *Phenomenology and the Cognitive Sciences,* 13 (3) : 461-475.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0004

Spread Mind and Causal Theories of Content

Krystyna Bielecka
Institute of Philosophy, University of Warsaw
kikab[]gazeta.pl
Received December 2013; accepted September 2014; published Autumn 2014.

Abstract

In this paper, I analyze a type of externalist enactivism defended by Riccardo Manzotti. Such radical versions of enactivism are gaining more attention, especially in cognitive science and cognitive robotics. They are radical in that their notion of representation is purely referential, and content is conflated with reference. Manzotti follows in the footsteps of early causal theories of reference that had long been shown to be inadequate. It is commonly known that radical versions of externalism may lead to difficulties with the notion of representation, especially if they cannot help themselves with the notion of syntax. I argue that a type of externalism present in Manzotti's enactivism may well lead to anti-representationalism.

Keywords: anti-representationalism; enactivism; causal spread; content; presentation.

In this paper, I analyze a kind of externalist enactivism defended, for example, by Riccardo Manzotti. Such radical versions of enactivism are gaining more attention, especially in cognitive science. However, it is not always appreciated that they involve radically externalist approaches to meaning. They are radical in that their notion of representation is purely referential, and content is conflated with reference. It is commonly known that radical versions of externalism may lead to difficulties with the notion of representation, especially if they cannot help themselves with the notion of syntax (or syntactic form, as in (Fodor 2008)). Without such notion, however, the enactivist view may be hardly tenable.

I argue that a type of externalism present in Manzotti's enactivism may well lead to anti-representationalism. Admittedly, he does not use the term "representation" and argues that the notion of "presentation" should replace it (Manzotti 2006). However, his "presentation" seems to play the role traditionally assigned to representation. Whether we call it "presentation" or "representation", the deeper problem is that it does not seem to answer the basic question of what its specific job would be, as opposed to any other factor that mediates perception and action. For this reason, his "presentational" view may collapse to anti-representationalism, especially in light of his arguments for mind's causal spread.

The aim of this paper is, first, to show some undesired consequences that such a kind externalism has for mental representation. Second, I want to show that even presentational or anti-representational views are in trouble when they help themselves with the notion of causation to define representation or presentation (or its anti-representational substitute). What is interesting, externalists such as Manzotti follow in the footsteps of early causal theories of reference that have long been shown to be inadequate. To justify this claim, I will show some similarities between radically externalist enactivism and early causal theories of Dretske (Dretske 1982) and Stampe (Stampe 1977). I will conclude by pointing out that some externalist theories, such as Manzotti's, are not acceptable because the problems they create are simply unsolvable.

1. Meaning and content

In this paper, I will appeal to two notions: meaning and content. Some initial explication is in order. These notions will be characterized as follows:

By "content" I will understand a non-formal property of a sign that allows us to distinguish two different expressions, even if they have the same reference. As a result, two expressions (e.g. "George Bush" and "ex-president") cannot be substituted in referentially opaque contexts, i.e.:

George Bush voted for Barack Obama during the last election in the USA.

The US ex-president voted for Barack Obama during the last election in the USA.

By meaning, I understand – just like Frege – reference (so a meaning of a "cat" is every instance of a cat, either black, white, brown, and so forth).

A theory of representation that does not include meaning or content as aspects of representation usually leads to inexorable difficulties. This is the reason why radical internalism (by eliminating meaning from the purview of the theory of mental representation) still remains so controversial. Today, the proponents of radical internalism have to offer additional reasons to justify their view against common objections (Segal 2000; Kriegel 2008) (I do not want to claim that it's impossible to defend radical internalism; all I imply is that this view simply requires further justification). Others either opt for two-factor theories (Block 1987; Block 1981) and juxtapose narrow content with a theory of reference or reject internalism altogether, as in the long-arm, or externalist, theory of representation (Greenberg and Harman 2006). While the latter tries to extend the notion of content so as to include the roles traditionally played by the notion of reference, the former tries to establish an appropriate relationship between the meaning and content of representations (which is not to say that this approach is entirely successful in this regard, see (Fodor and Lepore 1992)).

2. Manzotti's rejection of content

Before I proceed to present Manzotti's views on mental representation, some elucidation is in order. Riccardo Manzotti is mainly known for his work on consciousness, and he defends the view that phenomenal consciousness is also extended in space. In that, he goes further than (Clark and Chalmers 1998), who believe that only the cognitive mind is extended, not phenomenal consciousness. Yet Manzotti's view is systematic enough to imply a fairly clear account of representation (or "presentation" as he calls it). Also, his account of enactivism and externalism seems to be gaining popularity among roboticists and cognitive scientists (Manzotti's being roboticist himself). For this reason, it may be instructive to analyze the conceptual framework implied by a growing body of cognitive research. Here, I will focus only on his views on "presentation", putting aside his ontological or metaphysical position, as it is quite complex. I do not think that anything really important hinges on the fact that I do not analyze his process metaphysics at any length here. By avoiding it, I will simply deal with methodological assumptions that he shares with other enactivists, who consider the world to be its best representation (Brooks 2003). It will be instructive to see what the cash value of this claim is, regardless of metaphysical views of its proponents.

As a radical externalist, Manzotti rejects content in his account of representation. In this sense, his view is similarly extreme as radical internalism, which rejects meaning but not content. Manzotti's view is, in other words, purely referential. There are other proponents of purely referential accounts of representation, one of the most prominent being (Fodor 2008). The most important difference is that in contrast to Fodor, Manzotti does not seem to find any use for the notion of the syntactic form to explicate the why "George W. Bush" and "The US ex-president", even if co-referential, are not substitutable in referentially opaque contexts, whereas Fodor does. Also, Fodor has long rejected the simplistic version of the causal theory of content that Manzotti seems to presuppose.

Most naturalistic theories of representation find themselves in trouble when they try to reduce representational relations to relationships of similarity or covariance (Fodor 1984). Manzotti supports the covariance theory in its causal version even if he rejects the Cartesian gulf between the subject and the object. He argues that on the latter view, the emergence of representation seems to be a miracle:

> Up to now, the nature of the [representational – K.B.] relation is a tantalizing mystery. Different solutions have been proposed: correlation, causation, law-like causation, emergence, identity, supervenience. None has proved to be completely satisfactory. (Manzotti 2006: 47)

Manzotti insists that instead of "representation" we should say "presentation". He claims that representation

> (…) is used as an explanatory notion with a meaning of its own — a representation is something that presents (or re-presents) something else. (Manzotti 2003: 289)

He understands mind as spread physically and spatiotemporally beyond the skin. He illustrates this claim by using the metaphor of a rainbow:

> As a unity, although constituted by a series of physical drops of water in space reflecting the light in a certain way, cannot be defined without knowing where and how it will be seen. (Manzotti 2006: 50)

For this reason:

> (…) drops of water reflect the sunlight in the same manner, yet only those which have a particular geometrical relation to the observer, due to his/her position and to the direction of the sun rays, are seen as part of the rainbow (…) A given rainbow exists only when the observer is in a given position with respect to the external stimulus. (Manzotti 2006: 50)

In his view, there is no separation between mind and the world because what a particular mind perceives is in the continuous process common for particular minds and the world. What is more (and more worrying!), everything that an individual mind can perceive is veridical. He believes that everything we have in memory comes from the world:

> According to the process view presented here, memory and mental imagery have a phenomenal aspect because they are "perception delayed in time". Whenever we remember something, an uninterrupted causal chain originated in an object/event/state of affairs reaches its end in the brain. (Manzotti 2006: 65)

There are researchers who also think that memory does not require representation, so he is not alone here; however, they do not seem to claim that memory should be identified with perception (see (Freeman 1991) on the olfactory memory of rabbits).

However, Manzotti needs to say what illusion is (or at least seems to be if all perception is veridical indeed):

> I propose to see illusions as instances of infrequent correlations among physical events. Slightly more formally, I propose to see illusions as situations where an event C — normally perceived in conjunction with some other event A — is exceptionally perceived in conjunction with some other event B. (Manzotti 2006: 68)

And then he concludes:

> It is supposed that A is perceived instead of B. However what is perceived is, as in normal situations, C. The hypothesis of illusions is unnecessary and the continuity with the external physical world is maintained. (Manzotti 2006: 68)

But the claim that one can explain away misrepresentation in such terms is deeply problematic. Explaining illusions, or any malfunction, in terms of frequency is controversial. Take a simple biological example: sperm. Very few sperm actually serve their purpose (Millikan 1984: 29). Similarly, beavers may splash their tails much more frequently than there is real danger, but it is adaptive to splash your tail if you are a beaver. Infrequency cannot define what illusion is; to see that it is enough to realize that what people perceive as movies, are actually still images displayed quickly one. If Manzotti was right about illusions, then the illusion of movement that we experience when watching a movie would go after a sufficiently long exposure to it. However, movie lovers do not seem to experience disillusionment that there is no movement even if they spend whole days in the theater.

Even if we were to accept the view that illusions are just special cases of veridical perception caused by infrequent correlations, how could we check which cases are more correct and why? Imagine an almost blind person. His or her visual perception functions incorrectly almost all the time, so his or her perception is almost never correlated with the world events. But how does his or her case differ from the case of people with normal vision that see a stick in the water as bent? Classifying both cases as illusions would be a mistake. We would rather say that the second case is a common perceptual illusion while the first is a physiological inability to perceive objects. And how does it differ from cases when I mistake a roe deer in a forest to be a running horse? Memory illusions, in which a person reports of a past event that seriously deviates from the event's actual occurrence, seem to be problematic in this view as well (Roediger 1996). They may be as frequent as one wishes.

Are non-perceptual mistakes possible in Manzotti's view? It is hard to see how any non-perceptual mistakes could exist according to his view if all representation (or rather presentation) is based on perception. Manzotti is in pains to stress that there is continuity between mental presentation and perception. But some of our presentations do not seem to veridical, even if perception is not affected. Take the Capgras delusion (Hirstein and Ramachandran 1997). The Capgras delusion appears when a person consciously recognizes faces (this presentation is veridical) but because of the malfunction of the limbic system, there is a lack of an appropriate emotional response. As a result, a person looking at a face of her or his family member thinks she or he is looking at a doppelganger. Now, perceptual processes are fine, but we have a complex process that causes people to deny that they recognize their family members as family members. They recognize them only as doppelgangers of family members. Frequency has nothing to do with it.

As a consequence, the view that presentation is continuous with perception allows for no simple falsity in presentations. Even having empty concepts, such as Pegasus, seems to be an inexplicable anomaly for this theory. Just like Dretske had to say that there is no such thing as false information (Dretske 1982), Manzotti is pressed to deny false presentations. But then what is the cash value of such a notion of presentation? While Dretske appreciates the fact that agents do misrepresent (Dretske 1986) and claims that this representation is not reducible to semantic information (or natural meaning), Manzotti simply bites the bullet and says that there is no such thing as misrepresentation. But what explanatory use could his notion of presentation have? Definitely not the one that is traditionally associated with the notion of mental representation.

Manzotti defines the relationship between the mind and the world in terms of a causal network; but causal factors seem to be apprehended in a holistic fashion:

> If causal properties are not located on objects but depend on the causal network as a whole, then they cannot be located on a particular object. In a similar way, instead of being atomistically separate and autonomous, causes and effects become different ways of looking at processes. (Manzotti 2006: 54)

On Manzotti's view, there is no possibility to distinguish some causes as *the* causes of this particular presentation. This is so because of two reasons. First, it is difficult to determine the meaning of presentation in a continuous, never-ending process; second, everything is always relative to the particular mind. I will show the detrimental consequences for understanding presentation in these terms in detail below.

3. The disjunction problem

Imagine a person who calls water “vodka” by mistake. What is an extension of his or her term “water”? A priori, we have three possibilities:

a) “Water” denotes water

b) “Water” denotes water or vodka (water v vodka)

c) “Water” denotes vodka

Intuitively, only the answer (a) is correct. But a causal theory of reference does not give us any criteria necessary to fix the denotation in this standard manner. It cannot answer our question of what still does belong to the extension of the term T and what a new thing that does not belong to the extension of T is. This is where the so-called disjunction problem appears. There is no possibility to distinguish expanding the extension of the term from a mistake in using it. Actually, the causal theories, if they rely on the covariation of the use of the term T and the occurrence of referents (the simplified version of which is Manzotti’s frequency view), have to embrace the answer (b). The extension always grows, and it’s impossible to shrink it again. As the set of referents grows, our description of reference needs to include more and more disjunctions. Thus, the disjunction problem is an essential problem for causal theories as well as any theories that explicate representation in purely causal terms (Fodor 1984).

The informational semantics defended by Dretske (Dretske 1982) and Stampe (Stampe 1977) assumed that representation is reducible to information and that there is no such thing as false information. But, it means that there is no misrepresentation, and this is hardly credible. For this reason, informational semantics has been criticized as deeply flawed (Godfrey-Smith 1989). Manzotti’s claim that presentation is basically reducible to veridical perception leads to exactly the same problem. Dretske later even admitted that mere causal relations are not enough to explicate the notion of representation that are applicable to misrepresentation (Dretske 1986). Manzotti, however, cannot escape the problem at all. All he can do is deny that there is misrepresentation.

4. Anti-representationalism and misrepresentation

It is notable that the same arguments to effect that the mind is a spread entity are also used by anti-representationalists in order to justify their view (e.g., Paco Calvo, who even uses the same term “causal spread” that was introduced earlier by Andy Clark and Mike Wheeler (Garzon 2008; Clark and Wheeler 1998)). It does not take much to see that Manzotti’s theory of “presentation" is rather anti-representational or at least too weak to defend a concept of representation.

Anti-representationalists such as Calvo stress the continuity between the world and the mind; according to them, there is no difference between on-line and off-line representations. On-line representations appear only in the immediate presence of an object perceived, while off-line ones are present when the object is absent. The latter are taken to be paradigmatic cases of representations. By taking continuity seriously, we can easily show that there is always some link between the representational state and the previous state. The causal chain is actually never broken. Then, however, it's hard to distinguish the off-line representations from the on-line ones. It appears as if the presence of representations is always immediate. As the same kind of continuity is presupposed by Manzotti, it is not so hard to see that his notion of "presentation" has to be quite weak. It is exactly the same kind of entity as Clark's "on-line representation." But, as Calvo argues, one can explain phenomena with on-line representations by recourse to real-time dynamics, and in his view, this kind of theoretical entity seems to play no specific job assigned for representation at all. Let me elaborate on this.

In his recent book, Ramsey argued that tracking theories of representation (that take indicators to be paradigmatic cases of it) have difficulties with specifying what is specifically representational about them (Ramsey 2007). The idea that they influence behavior seems clear, but gravity does as well. The danger is that such theories may be too thin, having no role for content at all (note that I do not embrace Ramsey's argument that Dretske's or Millikan's teleosemantics is trivialized for this reason, as there is a role for content in their theories; but this is a topic for another discussion). All in all, dispensing with content and playing with continuity may be detrimental.

Interestingly, Fodor has recently also denounced any role for content, taking a purely referential view (Fodor 2008). However, there is a substitute for content in his theory. The form of the representation seems to differentiate "the US ex-president" and "George W. Bush" quite clearly. Just because Fodor can explain referential opacity by appeal to the form, his view does not suffer from the same objections that are detrimental to other referential views on representation. However, note that it is not possible for Manzotti to endorse such a view at all. He sees no role for vehicles of presentation (Manzotti 2006: 60) and tries to make them as dynamic as possible. It seems that the form of presentation would also be quite spread in time and space. Manzotti, for this reason, appears to think that the notion of vehicle and its form is dispensable (note that this is not in general accepted by all proponents of the dynamical view on cognition; see e.g. (Rączaszek-Leonardi 2009; Deacon 2011)). But then, referential opacity remains an inexplicable mystery for his theory.

Moreover, what lacks in Manzotti's theory is a satisfactory concept of misrepresentation. Even if we agreed that misrepresentation is an on-line representation located in a continuous process, his account of misrepresentation in terms of frequency would be totally unsatisfactory. Illusions, delusions, hallucinations, or simple fallacies in reasoning can appear very often and some of them seem to be innate (e.g. gambler's fallacy).

It's worth noticing that anti-representationalism also cannot avoid describing these phenomena and their description in terms of a continuous causal chain or process is not informative. Understanding representation in terms of spatiotemporally spread chain seems to be drastically incomplete. Representational theories do not state that representation is just veridical perception and it is as if anti-representationalists try to avoid explaining what else should be added but that perception will help describe fallacies or delusions.

Without a satisfactory notion of misrepresentation, no account of representation is acceptable (Dretske 1986). But this is also true of accounts of cognition, anti-representational included. A representational theory that denies the very possibility of there being any misrepresentation is a theory that deflates the concept of representation and trivializes it completely. Such representations may be indeed epiphenomenal. If you want to explain my behavior only by recourse to true beliefs, then I wish you good luck, especially because some of these beliefs are not mine at all.

5. Conclusion

In this paper, I argued that some radical externalist theories, such as Manzotti's theory of "presentations", are equally misconceived as early causal theories of reference defended by Dretske or Stampe. I want to draw two general conclusions from this fact. These conclusions are pertinent to radical externalist theories such as Manzotti's as well as to theories of representation in general, enactivist or not.

First, my criticism of Manzotti's theory may apply to any radical enactivist theories that are anti-representational or as weakly representational as Manzotti's. The similarity between Manzotti's enactivism and the early causal theories of reference shows that they both deflate representation. The ability to tackle misrepresentation is what should be required from any theory of representation that is psychologically plausible.

Second, it is remarkable that theories of representation that reject either content or meaning have difficulties with explaining what job the representation has *qua representation* in the cognitive system. To reject meaning (as radical internalists do) is to make it difficult to understand the relationship of representations and the world in which the cognitive system finds itself in. The theories of this kind are therefore solipsistic. But to reject content (as radical externalists do) is to make misrepresentations and false representations impossible, although the problem with relation to the world does not arise.

Note that Manzotti's troubles with misrepresentation and content are not at all implied simply by his enactivism. They arise because of a grossly simplified view on "presentation" that seems to take the claim that the world is its best presentation seriously. If this is so, then presentation cannot be wrong. However, this also means that there cannot be learning, as learning implies the ability to correct previous mistakes and recognize them as such. If enactive approach is the one that recognizes the crucial inter-dependency between the autonomous agent and the world it finds itself in, then many theories of representation seem to be compatible with it (and some of them deal explicitly with the disjunction problem or misrepresentation; see (Bickhard 2008; Bickhard 1993; Anderson and Rosenberg 2008)). The trouble is that it is all too easy to radically deny the need for representation. I dare say that this means denying the need for cognition and learning as well, but that is a matter for further discussion.

References

Anderson, M.L., and Gregg Rosenberg. 2008. Content and Action: The Guidance Theory of Representation. *Journal of Mind and Behavior* 29 (1-2): 55–86.

Bickhard, M.H. 1993. Representational Content in Humans and Machines. *Journal of Experimental & Theoretical Artificial Intelligence,* 5 (4) (October 1): 285–333. doi:10.1080/09528139308953775.

Bickhard, M.H. 2008. The Interactivist Model. *Synthese,* 166 (3): 547–591. doi:10.1007/s11229-008-9375-x.

Block, N. 1981. Psychologism and Behaviorism. *The Philosophical Review,* 90 (1) (January 1): 5–43.

Block, N. 1987. Advertisement for a Semantics for Psychology. *Midwest Studies in Philosophy,* 10 (1) (September): 615–678. doi:10.1111/j.1475-4975.1987.tb00558.x.

Brooks, R. A. 2003. Intelligence with Representation. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences,* 361 (1811) (October): 2381–95. doi:10.1098/rsta.2003.1257.

Clark, A., and D. J. Chalmers. 1998. The Extended Mind. *Analysis,* 58 (1): 7–19.

Clark, Andy, and Michael Wheeler. 1998. Bringing Representation back to Life. *From Animals to Animats 5-Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*. R. Pfeifer, B. Blumberg, J.-A. Meyer, and S. W. Wilson, eds. 3–12. Cambridge, Mass.: MIT Press.

Deacon, T.W. 2011. The Symbol Concept. *Oxford Handbook of Language Evolution*, edited by M. Tallerman and K. Gibson. Oxford.

Dretske, F. I. 1982. *Knowledge and the Flow of Information*. 2nd ed. Cambridge, Mass.: MIT Press.

Dretske, F. I. 1986. Misrepresentation. *Belief: Form, Content, and Function*. R. Bogdan, ed. 17–37. Oxford: Clarendon Press.

Fodor, J. A. 1984. Semantics, Wisconsin Style. *Synthese* 59 (3) (June): 231–250. doi:10.1007/BF00869335.

Fodor, J. A. 2008. *LOT 2: The Language of Thought Revisited*. New York: Oxford University Press.

Fodor, J. A., and E. Lepore. 1992. Why Meaning (Probably) Isn't Conceptual Role. *Mind & Language*, 6 (4): 328–343.

Freeman, W. J. 1991. The Physiology of Perception. *Scientific American,* 264 (2): 78–85.

Garzon, F. C. 2008. Towards a General Theory of Antirepresentationalism. *The British Journal for the Philosophy of Science,* 59 (3) (September 1): 259–292. doi:10.1093/bjps/axl007.

Godfrey-Smith, Peter. 1989. Misinformation. *Canadian Journal of Philosophy,* 19 (4): 533–550.

Greenberg, M., and G. H. Harman. 2006. Conceptual Role Semantics. *The Oxford Handbook of Philosophy of Language*. E. Lepore and B. Smith, eds. 295–322. Oxford: Oxford University Press.

Hirstein, W., and V. S. Ramachandran. 1997. Capgras Syndrome: A Novel Probe for Understanding the Neural Representation of the Identity and Familiarity of Persons. *Proceedings: Biological Sciences,* 264 (1380) (March 22): 437–444. doi: 10.2307/50434.

Kriegel, U. 2008. Real Narrow Content. *Mind & Language,* 23 (3) (June): 304–328. doi:10.1111/j.1468-0017.2008.00345.x.

Manzotti, R. 2003. A Process-Based Architecture for an Artificial Conscious Being. *Process Theories: Crossdisciplinary Studies in Dynamic Categories*. Johanna Seibt, ed. 285–312. Dordrecht: Kluwer Academic Publishers.

Manzotti, R. 2006. An Alternative View of Conscious Perception. *Journal of Consciousness Studies,* 13 (6): 45–79.

Millikan, R.G. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, Mass.: The MIT Press.

Ramsey, W.M. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511597954.

Rączaszek-Leonardi, J. 2009. Symbols as Constraints: The Structuring Role of Dynamics and Self-Organization in Natural Language. *Pragmatics & Cognition,* 17 (3): 653–676. doi:10.1075/p&c.17.3.09ras.

Roediger, H.L. 1996. Memory Illusions, 100 (35): 76–100.

Segal, G. 2000. *A Slim Book about Narrow Content*. Cambridge Mass.: MIT Press.

Stampe, D.W. 1977. Toward a Causal Theory of Representation. *Midwest Studies in Philosophy,* 2 (1) (September): 42–63. doi:10.1111/j.1475-4975.1977.tb00027.x.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0005

Agency is Distinct from Autonomy

Fred Cummins
UCD School of Computer Science and Informatics
University College Dublin
fred.cummins[]ucd.ie
Received December 2013; accepted June 2014; published Autumn 2014.

Abstract

Both autonomy and agency play central roles in the emerging enactive vocabulary. Although some treat these concepts as practically synonymous, others have sought to be more explicit about the conditions required for agency over and above autonomy. I attempt to be self-conscious about the role of the observer (or scientist) in such discussions, and emphasise that the concept of agency, in particular, is deeply entwined with the nature of the observer and the framing of the observation. This is probably well known to enactivists, but runs the risk of being badly misunderstood if it is not made explicit. A heightened awareness of the role of the observer in the attribution of agency may allow us to make advances in questions in which progress is hindered by assuming a single split between subject and object. I argue that human experience is characterized by our embedding in webs of meaning arising from our participation in systems of many sorts, and that this richness demands a corresponding lightness of touch with respect to the identification of agentive subjects.

Keywords: agency; autonomy; enactivism; cognition; mind.

1. Introduction

The treatment of agency in human intercourse is a matter of the highest importance. It underpins the enlightenment notion of human rights. It informs, or should inform, the manner in which society apportions responsibility for actions, including those boundaries that delimit the criminal and the insane. The attribution of agency is one of the principal themes that informs and defines all religious traditions. When the enactive and mind-and-life schools choose to make agency a foundational concept, they must do so with a sense of gravitas, for the issues at stake go beyond novelty in the cognitive and behavioral sciences or philosophy of mind. In what follows, I will consider the entanglement of the twin concepts of autonomy and agency as they feature in the enactive literature. At stake is whether these terms pick out distinct concepts, and more importantly, whether agency can and should be naturalised, and what that would mean.

I begin by acknowledging the consequences of a constructivist approach for the practice of science, which demands that we recognize the social and historical context in which the attribition of agency has been restricted to a single kind of split between agentive subject and associated world. Some careful attempts to characterize agency over and above autonomy are discussed in detail, leading to the cautious stance that an overly rigid attribution of agency to this or that system may blind us to the important distinction between the domain of the system and the domain of the description of the system. This caution allows us to recognize the plurality of systems among which our lives unfold, that provide many and varied examples of systems that may appear agentive. The hope is that this caution will allow us to recognize that in attributing agency here, or there, we are at the same time identifying our own being with diverse forms of systematic organisation and hence with diverse systems of values that thereby arise.

The enactive agenda necessarily has one foot in scientific practice and another in the philosophy of science (there may be more than 2 legs). As Stewart noted "the paradigm of enaction is ontologically nonobjectivist—or to put it more positively, radically constructivist" (Stewart et al. 2010). Constructivist agendas have not found easy integration into mainstream scientific practice, and indeed many, perhaps most, areas of scientific inquiry get by just fine with the presumption that they are uncovering the structures and form of a mind-independent world. For those of a constructivist bent, this is, of course, no longer tenable for "sciences of the mind", or, by extension, for any of the Human Sciences.

A necessary consequence of adopting a constructivist perspective is that explanation is necessarily bounded. In the spirit of pragmatic explanation broadly construed, the practice of explanation, description and modeling is not to fix this or that phenomenon within a single static account of the world, but to develop an understanding that is appropriate to a specific domain, and that may serve in a broader array of accounts, some more specific, some more general, but none aspiring to be ground truth. The pragmatic boundedness of explanation here is neither the logical kind espoused by Pierce (Murphy and Rorty 1990) nor the unbridled catagorical cornucopia of Ryle (Ryle 1949), but demands that explanation be couched in terms appropriate to specific domains. The biological concept of autopoeisis identifies a particularly important form of autonomy that is defined with respect to the chemical domain (Maturana and Varela 1991). A more general concept of autonomy, in Varela's definition, requires inter alia that the processes that characterize an autonomous system be a unity "recognizable in the space (domain) in which the processes exist" (Varela 1979: 55). In an informative example, Varela suggests that "the accidental collision of two running animals, as a bodily encounter of living systems, is not a biological phenomenon (even though it may have biological consequences), but the bodily contact of two animals in courtship is." (Varela 1979: 42).

Most scientific practice is not done within a constructivist framework, and an appreciation of the limited nature of explanation, of the finiteness of domains of discourse, is not something that can be taken for granted among scientists generally, and still less so in the conversation between science and society more generally, where science is still almost universally regarded as the source of certainty, and the ultimate arbiter of disputes. As the enactive agenda develops, and seeks to provide a set of concepts applicable in domains from the biochemical to the social, it must exercise caution if it is not to be enthusiastically misunderstood as a means of establishing certainty with respect to many vexed notions that it takes as central. Among these are the twin terms of autonomy and agency. This is a first challenge we face.

In its principled rejection of the Cartesian split of things into {inner, mental, and subjective}, versus {outer, material, and objective} realities, enactive theory finds itself in the lexically challenged business of talking about human experience and behavior without being able to lean confidently upon any psychological predicates whatsoever. Yet we cannot do without these. To address this, it is necessary to recognize a second consequence of a radical constructivist approach to understanding: As we learn about phenomena in various domains, so too we are learning about that which we are.

> Maxim: Along with asking "what is this thing that we see", we need to ask "what are we that we should see such? ".

A corrolary of this is that our understanding of *all* psychological predicates should be taken as tentative, and subject to change.

The emphasis in that maxim is on the "we". The sociocultural background in which scientific practice has developed has the striking characteristic of attributing agency in human affairs exclusively to the individual person, conceived of as a discrete organic unit. Democratic societies that emerged after the Age of Revolutions are founded upon the notion of individual agency. The prevalent Christian theology, and post-Reformation Protestant ideology position the concepts of culpability and responsibility squarely within the individual. It is against this backdrop that the science and practice of psychology emerged in the latter half of the Nineteenth Century, and the unit of psychology, even social psychology, is the individual person, conceived of as a singular mind housed within a singular body.

The emerging enactivist position has the potential to move our discourse beyond the limitations of Protestant theology and 19th Century psychology. The language that is emerging, and being carefully tended and curated, allows explanation to countenance multiple perspectives, even mutually incompatible perspectives. In its parallel treatment of system-internal and system-external perspectives, it allows for a richer, plural, understanding of the interactions among complex systems such as cells, persons, or social groups, and a recognition of disparate value sets that influence their evolution and interaction. But a positivist view of scientific understanding and a restriction of agency to single human individuals are both deeply entrenched in our ambient belief systems and practices. Careful exegesis of the conceptual fulcrums of the enactive approach is indispensable, and the treatment of agency in particular will be of singular importance. It would be unwise to be dogmatic here.

2. Pinning Agency to the Wall

Autonomy and agency are absolutely central to the enactive approach.[39] The somewhat vague notion of autonomy was given a more precise definition by Varela as "organizational closure" (Varela 1979: 58). On this view, autonomous organization is characterized by circular closure among a suite of processes that collectively constitute a persistent dynamic identity that engages in regulated exchanges with its surround. The most discussed example, by a substantial margin, is the caricature of a single bacterium ascending a chemical gradient through chemotaxis. This example, trotted out repeatedly with differing degrees of attention to biochemical detail, serves as the exemplary embodiment of the autonomous and agentive system. As an illustrative case, it allows discussion of value, as the nutrient is unambiguously a "good thing" from the point of view of the bacterium itself. It allows discussion of sense-making by linking the effective coping of the bacterium in a variable environment with its own metabolic requirements. It illustrates the contrast between a perspective anchored to a specific system, and a view from nowhere. In short, the chemotactic bacterium serves as the central myth of the enactive approach, in the sense of a narrative that serves to structure many discussions that need to be had as we apply systems thinking to our own selves.

But real bacteria are vastly different, more complex, more tightly embedded in their environments, and more social than this. The bacterium of this oft-repeated illustration is described as a minimal *mechanism*, requiring a single sensor capable of detecting ambient glucose concentration, a means of locomotion with a directed mode and an undirected mode, and a probabilistic link between the slope of the ambient gradient and the likelihood of switching

[39] I do not make any strong distinction between enactive theory and mind and life positions within philosophy. Both terms cover a range of positions, with very substantial overlap.

between the two modes. Viewed as a mechanism, it is indistinguishable from any machine concocted by a designer, and indeed, as a character in our narrative, it could be argued that it has indeed been designed. There is nothing to be found in the mechanism that warrants the attribution of agency. That attribution comes instead from our understanding, as scientists, that chemotactic locomotion can be understood as serving the continued existence of the cell as an organisationally closed set of processes. A functional description, by us as scientists, is predicated upon our passing adoption of the perspective of the bacterium, which is the domain for whom the ambient glucose is meaningful, thereby licensing talk of "function".

In discussion of the caricature of the bacterium, two related issues seem to become entangled. One issue concerns the difficult task of differentiating between machines that are designed, and machines that have evolved. The second lies in distinguishing between "merely" autonomous phenomena such as tornados and flames on the one hand and apparently agentive (and autonomous) cells on the other. With respect to the first distinction among machines, there is no overt marker of the locus of the designer, or the origin of the organising principle that characterises the machine. Thus that distinction is made on the basis of our knowledge of the history of the machine. With respect to the second, there is much discussion that remains to be had, but the issues at stake are importantly different.

An informative example of the confusion of the concepts of autonomy and agency is provided by two related papers. In Rohde and Stewart (2008), it is argued that *autonomy* is a complex property, similar in some respects to such complex notions as intentionality or intelligence. When faced with a system (natural or artificial), the ascription of autonomy thereto can not be based on a set of necessary or sufficient conditions. But the authors are not happy with the "as if" ascription that falls out of adopting what Dennett has called the "intentional stance" with respect to the system (Dennett 1989). Taking a constructivist perspective, the authors seek to improve on mere ascription through appeal to generative mechanisms that can give rise to the phenomenon of autonomy. Taking a cue from that suggestion, then, Barandiaran et al. (2009) attempt to define the concept of *agency*, first through mere description, and then improving and deepening that characterization through the proposal of a set of generative mechanisms that can give rise to agency. The switch from a generative description of autonomy to a generative description of agency provides a telling example of the manner in which these two concepts have become entangled. Note that the explanatory move common to both these papers is independent of the distinction I am drawing here. Both papers seek to ground an account of a complex phenomenon in something more than mere description, and both do so by appealing to generative mechanisms. Whether this succeeds or not is not at issue here. What is at issue is whether autonomy and agency are separate dissociable concepts. Let us look at some examples from the recent literature:

Moreno and Etxeberria (2005) contrast the “actions" performed by a stone in a river that keeps water from flowing, and by bacteria in milk that ferment it. Only the latter, they say, can lay claim to agentive goings on, and they ground this claim in the role of the fermentation from the perspective of the bacterium. This contrast seems particularly clear, but mainly because the rock is not “doing" anything. Absent activity on the part of the agent, agency is invisible. Barandiaran et al. (2009) consider a suite of more interesting and challenging contrasts that provide guidance in the application of three criteria: *Individuality* (is the system an individual in the above sense?), *Interactional Asymmetry* (is the system the active source of interaction with its environment?), and *Normativity* (is the norm that shapes the interaction generated by the system itself?). Simultaneous satisfaction of these three conditions serve to pick out the chemotactically locomoting cell as the only instance of full agency, while several other carefully chosen candidates fail to tick at least one of the three boxes. Individuality corresponds roughly to the definition of autonomy as organisational closure, as discussed above. Both cells and tornados might reasonably meet the criterion of systemic individuality. We therefore need to consider whether the other two conditions succeed in going beyond autonomy and to pick out agency as a distinct and definite concept.

The requirement of normativity refers to the interpretation of a behaviour as subserving goals, and Barandiaran et al. (2009) wish to restrict the attribution of agency to systems which generate their own goals. Many authors have recognized that agentive behaviour occurs in the service of goals, indeed its goal directed nature may be the very feature that allows a continuous stream of movement to be parsed by an observer into discrete behaviours for consideration in the first place. Some have opined that goals alone are sufficient for agency (Beer 1995), but most researchers, especially those who seek to distinguish the agency of the living from the functional carry-on of devices and robots, have insisted that agency requires that the goals arise from the needs and identity of the system itself (Weber and Varela 2002; Maes 1993; Christensen and Hooker 2000; Kauffman 2002; Deacon 2011, and others). Upon this rock, many have foundered, even Immanuel Kant himself, for whom the intrinsic purposes of the living posed an insurmountable challenge within a Newtonian metaphysical view. Unlike Kant, we have available to us a richer metaphysical armoury and the powerful set of concepts that stem from the description and study of complex systems (Weber and Varela 2002). However, even with this endowment, there is not available to us any litmus-test for distinguishing between intrinsically generated and extrinsically imposed goals in the shaping of behaviour.

A first example adduced by Barandiaran et al. (2002) meets the criterion of intrinsic normativity, but is still regarded as non-agentive. Consider a mother cat and several kittens. One of the kittens is moved by the mother closer to her body to keep it warm. The kitten meets the criterion of dynamic identity, and it also satisfies the intrinsic normativity criterion because “the system-environment coupling is satisfying the norm of keeping the kitten’s temperature within viability boundaries" (Barandiaran et al. 2009: 5). Agency is lacking, the authors contend, because the movement has its origin in the mother—a separate system that forms part of the kitten’s environment. This example is threatened by the simple expedient of re-bracketing the system under observation: if the system is the family of cats, the source of the action now lies within the bounds of the system, the identity and normativity conditions are met, and it seems that the act is now agentive in the strong sense the authors seek. I do not wish to argue that the authors are correct or incorrect in the attribution of agency. I wish to point out that the coherence of the notion depends upon the framing of the observation. To the extent that the family can be considered a system with its own dynamic identity, it is available as a possible locus of agency[40]. One of the core insights of the enactive approach is that autonomous organization is not fixed at one level. In mutual interaction, pluralities of autonomous systems can form novel superordinate emergent domains that themselves exhibit autonomy. With autonomy arises the need to consider the perspective of the emergent autonomous system, and the phenomenology enacted thereby. This general approach can be taken in considering interactions among cells, producing multicellular phenomena, including multicellular organisms. It can be applied to the emergence of social phenomena, e.g. riots or Mexican waves, when people engage in rich, reciprocal interaction with each other and within the confines of just the right set of constraints. And it can be applied to a mother cat and her kittens as a whole.

Related objections arise in the case of Parkinsonian tremors. Here, it is claimed, the human body is the system to be considered, the system is the active source of the interaction, but the normativity condition is not met as the tremor does not serve to maintain, protect, or establish any internally generated norm. Elsewhere, the argument is made that “the spasms of a person from Parkinson’s disease are not considered to be the [sic] actions, even though the person is a well-identifiable entity and the genuine source of her interactions with the environment” (Barandiaran et al. 2009: 5). The system has been variously identified here as the body, and the person. Neither seems entirely appropriate. Tremor arises when the equilibrium dynamics of a somatic subsystem changes from a static equilibrium to an oscillatory regime.

[40] Perhaps the mother + kittens is a poor candidate for the ascription of autonomy in the first place. In that case, my argument in this instance is weakened, but the general observations about the role of the framing of the system remains.

The fact that an oscillatory dynamic arises at all is evidence that there is indeed a system underlying the phenomenon, but it is not the whole body. The persistent copulatory motions of a male preying mantis after decapitation illustrates the notion of an oscillatory regime in a system that is less than a whole body rather well.[41] The assignation of the tremor to the body is thus misplaced, but the confusion of the body and the person is a far more serious matter, to which we will return in the next section.

The final example adduced by Barandiaran et al. (2002) is that of passive osmosis at the boundary of the cell. Here, the system seems to be uncontroversially the cell, and if the osmotic process subserves metabolism, the normativity condition is met. Agency, it is claimed, is not present because the system (the cell) is not the "active" source of the interaction. This example seems to typify a deep linguistic problem when discussing agency. Many verbs of action implicitly lean upon the background notion of an actor. Thus if we say that a property *p* of a system *s* is modulated, it is hard to make sense of that statement unless there is an implied agent, the modulator, who is doing the modulation. The asymmetry condition employed in the definition of agency in Barandiaran et al. (2009) seems to become circular for this reason when they say:

> We therefore define interactional asymmetry as the condition describing a system as capable of engaging in some modulations of the coupling [between system and environment] and doing so at certain times. (Barandiaran et al. 2009: 4)

If a system "engages in some modulations", we might well view it as an agent, but the presumption of agency lies here in the very description that employs an agentive verb. Such circularity infests the literature on agency. Kauffman (2002: 49) defines an agent as a system doing something on its own behalf. The verb "doing" carries the implication of a do-er, and circularity threatens again.

3. The Eye of the Beholder

Both asymmetry and normativity, it seems, depend upon the framing of the phenomena by an observer. In this, they have a lot in common with the individuality criterion itself. As Rohde and Stewart noted, taking the status of the observer seriously "transforms our conceptual world in a way that blurs the boundaries of what we normally consider a belief and what we consider a fact" (Rohde and Stewart 2008: 425). This groundlessness that necessarily arises when one eschews a positivist commitment presents something of

[41] If truth be told, this example comes from a Tom Waits song, and not a detailed study of the entomological literature.

a challenge if claim and counterclaim are to be judged on criteria other than opinion. The pragmatic approach suggested in Rohde and Stewart (2008) for treating of autonomy *simpliciter* and followed by Barandarian et al. (2009) is to shore up the justification of one or other perspective by the postulation of generative mechanisms. In so doing, they seek to go beyond mere description, while stopping short of an inflexible essentialism. But the ability to characterize a generative mechanism, by itself, does not achieve what the authors seek. It is possible to provide algorithmic effective procedures to illustrate many kinds of process, but that does nothing to establish the accuracy of the mapping from the algorithm to the world. This is a rhetorical strategy that has, in the past, lent the domain of artificial intelligence and some varieties of cognitive psychology a veneer of objectivity without due warrant. For example, proponents of a very different kind of agenda have sought to justify the creation of elaborate representational mental models using the same argument:

> The theory should be describable in the form of an *effective procedure*…If a procedure can be carried out by a simple machine, plainly it does not require any decisions to be made on the basis of intuition or any other such 'magical' ingredient: it is an effective procedure. (Johnson-Laird 1983: 6)

By avoiding appeal to intuition, or even magic, the resulting theory is cast as objective, and hence inherently trustworthy. But the argument does not even try to ascertain the appropriateness, viridicality, or utility of the relation that obtains between the elements of the theory and the nature of the phenomenon being studied. So the appeal to generative mechanisms may serve to convince friends, but will hardly convert sceptics.

An alternative perspective on our strong predilection for attributing agency to one system but not another is provided by the claim that life can only be known by life. Because we are ourselves beings whose existence is a continual striving, we recognize this striving in others. If we were disembodied intellects, no such communion with the living would make sense, and the concepts of *organism* and *agency* would not have any grounding. Thompson argues:

> To make the link from matter to life and mind, from physics to biology, one needs concepts like *organism* and *autopoiesis*, but such concepts are available only to an embodied mind with firsthand experience of its own living body. (Thompson 2004: 90)

This argument has been frequently made, and is sometimes attributed to Hans Jonas (1968). It is unclear to me whether both of the concepts of autonomy and agency should be considered to be members of the set of concepts that can only be understood by an embodied living being. It does not seem to stretch credulity too much to assume that tornados and flames might be recognised as examples of dynamically individuated phenomena to an abstract or alien intelligence with the ability to make observations at a human-like time and spatial scale. Jonas dismisses their individuality on the grounds of intuition alone (Jonas 1968: 240), but Jonas is not drawing a distinction between the organisational closure and dynamical persistence that characterises autonomous organisation on the one hand, and the striving of an agent exhibiting behaviour regulated by goals or norms on the other.

There is thus a tension in the literature that informs the enactive perspective. On the one hand, Maturana (and possibly Varela) insists that:

> Purpose or aims …belong to the domain of our discourse about our actions, that is, they belong to the domain of descriptions, and when applied to a machine, or any system independent from us, they reflect our considering the machine or system in some encompassing context…Accordingly, if living systems are physical autopoietic machines, teleonomy becomes only an artifice of their description which does not reveal any feature of their organization, but which reveals the consistency in their operation within the domain of observation. Living systems, as physical autopoietic machines, are purposeless systems. (Maturana and Varela 1991: 85–86)

On the other, Jonas objects that biological organisms simply are not machines:

> [W]hen we call a living body a "metabolising system," we must include in the term that the system itself is wholly and continuously a result of its metabolising activity, and further that none of the "result" ceases to be an object of metabolism while it is also an agent of it. For this reason alone, it is inappropriate to liken the organism to a machine…food is more than fuel…its role is to build up originally and replace continually the very parts of the machine. Metabolism thus is the constant becoming of the machine itself—and this becoming itself is a performance of the machine: but for such performance there is no analogue in the world of machines. (Jonas 1966: 76)

Given the degree to which Varela leans upon the work of Jonas later (e.g. in Weber and Varela 2002), there does not seem to be a fundamental disagreement here, but rather an occasional failure to knowingly distinguish between the domain of description, as things appear to an observer who is, herself, an agent, and the domain of operation of the system. Recognizing this, it seems to me, opens up opportunities to further our understanding of autonomy and agency in new ways.

4. Webs of Meaning

As living entities, we are enmeshed in webs of significance and meaning. The tantalising prospect opened up by the enactive approach is to lay the foundation for a rational, scientifically informed account of our goings-on that can lean upon such notions as value and meaning, without descending into mere relativism and the brandishing of opinions. But this must be done self-consciously. When we speak of a value, it is with respect to some system for which we recognize a degree of autonomy. We can ground such discussion through observation of the degree to which a given system exhibits organisational closure, by examining the manner in which the system interacts with its immediate environment. We can do this to understand the conditions under which the autonomous identity exhibited by the system will persist, and the conditions under which it will be threatened.

The construct of the person and the organic reality of the human body need to be clearly distinguished here. Some behaviours exhibited by my movements make sense when referred to the domain of the body, as, e.g. when I recoil from a physical threat, seek water to still my thirst, or sleep. Very many of my behaviours do not make sense with respect to the somatic domain alone. When I go to a movie, play a tennis match, or vote in an election, these actions are structured by value, but the interpretation of the behaviour as subserving goals arising from one or other domain is not straightforward. There is no guarantee that an activity that can be picked out linguistically ("playing a game of tennis") can be interpreted with respect to the values arising from any single domain. If the game of tennis is a doubles match, we can recognize the domain of my body, the dyadic domain that is my team, and the set of four players who together enact the game of tennis. Some features of the activity are best understood with respect to each of these three domains, and there are probably further domains that one could sensibly identify. Furthermore, both the somatic individual, and the dyad that is a team, may admit of agentive interpretations, though I see no such interpretation for the set of four players.

Where an agentive ascription seems appropriate, the language of intentionality, of goal directedness and striving, can sensibly be applied. The desires, intentions and sensibilities of a group of protesters chanting in unison seem, to this author, to be no more or less real than the desires, intentions and sensibilities of me on the tennis court trying to win a match. Intentional predicates are easily and naturally used with respect to institutions, nation states, mobs, armies, and teams. On a conventional psychological reading, these uses are metaphorical extensions of the one true sense grounded in the (Cartesian, solipsistic) mind of the cherished individual. But from an enactive perspective, we can be flexible not only with respect to the identification of domains that exhibit temporally extended dynamic identity, but also with respect to the implied subject, that provides the values and normative scaffolding that allows mere movement to be interpreted as action or behaviour.

This constitutes something of a reversal of the normal course of scientific inquiry. The maxim introduced earlier encourages us to constantly ask both what do we see, and what is the implied subject of that seeing. This groundlessness is well known within the traditions of Buddhist and Taoist Philosophies, but has only rarely been acknowledged as a valid epistemological stance within the domain of science (Varela et al. 1991). In ascribing agency, we are tacitly acknowledging a commonality with the system being observed–

not a commonality based on the human body, but based on the lived experience of striving that is at the center of the experiencing subject.

5. Disinterest and Empathy

There is here an opportunity to develop a principled manner in which intentionality and value can be rationally treated within our collective discourse. Where the received psychological tradition insists on a single, immovable, split between subject and world, we are now in a position to recognize that any such split lies within the domain of description of this phenomenon or that. As our discourse ranges over the affairs of the body, of the family, of the schools and institutions of society, of the relation between man and climate systems, and beyond, the ground beneath the discussion shifts, and agency, attributed now this way and now that, brings into being first this subject and now that. Some of these subjects will appear singular, but some speak of collective intentionality, collective values, values originating with respect to an ever-changing topicalised object.

I bemoaned earlier the casual description of the domain underlying a Parkinsonian tremor, first as the body, and then as the person. Neither is the appropriate domain to ground the observation of the tremor, but the confusion of the body and the person is a far more serious confusion. We can speak with a degree of disinterest of the body. For many of us, we can contemplate replacing limbs, organs, and the like, with artificial prosthetics, without feeling that the person associated with the body is thereby fundamentally changed. For when we speak of the person, we speak of the value-laden ground of experience, of subject-hood. And the subject is neither singular nor plural. Fixing the subject at one level, as conventional psychological theory does, privileges some values over others, introduces a normativity that is utterly at odds with the richness and dynamic constitution of our being.

We can ascribe autonomy in a disinterested fashion, and we should do so. In this way we can distinguish between the chimera of the transient and the persistent identity of the autonomous system. When we go further and ascribe agency, we take sides, and express a degree of empathy, however slight. In the agentive description we acknowledge some relations between a system and its milieu as privileged, and accord them value. In dealing with the goings on of the animate, we have no choice but to traffic in the values of the animate. With life, value leaks in.

We must, therefore, be careful not to conflate the distinct notions of autonomy and agency. Autonomous organisation can be recognised in systems that do not display agency, as well as those that do. The ascription of autonomy in the sense of organisational closure or dynamic identity is done based on a set of criteria that are rooted in our observations. Changing the timescale, spatial scale, or granularity of our observation may lead us to recognize that any ascription of autonomy is context-dependent and is done self-consciously by us as observers. The ascription of agency goes much further, and leads to the recognition of value—not an unprincipled recognition of value, but one grounded in system identity. However as observers, we become more firmly entwined with the objects of our observation when we engage in discussion of norms, goals, and values. These discussions must be done with an even greater degree of care.

There are differing views about the future utility of the enactive perspective. In some respects, it does not sit neatly within the contingent administrative divisions that characterize our institutions of learning and research. It is not psychology, nor is it merely social science, nor biology, and heaven forfend, it should not be mere philosophy. There are very many issues that are well addressed within the conventional frameworks, and for those there is no urgency in insisting upon an alternative approach. A cardiologist dealing with an acute heart attack does not need to ponder whether the function of the heart is better referred to the domain of the body or the domain of the description of the body, or just what the domain of the subject here is. The rehabilitation worker who needs to address the difficult re-integration of the post-heart-attack patient into the overlapping and competing spheres of family and occupational life, on the other hand, might have need of some principles to guide that complex task.

The advantage to be gained from adopting a fundamentally new perspective is that questions may now be posed, and perhaps addressed, that lie beyond our present competences. It is therefore to be hoped that the careful curation of the enactive vocabulary may help us to untie the conceptual knots that arise from a sterile and immovable split between mind and world. Here the ill-structured theoretical quagmire of mental health appears as an obvious domain of potential application. But so too do numerous societal issues that bear upon the relation of the individual to the many forms of collectivities that together make up our lives. Thinking further ahead, the collective that is humanity will inevitably face major issues for which it is ill-prepared, and that demand working solutions couched in a vocabulary of a collective that is neither imposed nor fictitious. The “we" will be negotiated, and that is only possible if we are capable of an ontological light touch in the matter of the (many) subject(s).

6. Acknowledgements

Marek McGann contributed mightily to lengthy and ongoing discussions around these themes, and while he cannot be held responsible for anything herein, if it has merit, some of that lies with his contributions. The little-known interpreter of Buddhist and Taoist texts who wrote under the name of Wei Wu Wei also lurks unacknowledged in the undergrowth of this article.

References

Barandiaran, X. E., Di Paolo, E., and Rohde, M. 2009. Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5): 367–386.

Beer, R. D. 1995. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72(1): 173–215.

Christensen, W. and Hooker, C. 2000. Autonomy and the emergence of intelligence: Organised interactive construction. *Communication and Cognition-Artificial Intelligence*, 17(3-4): 133–157.

Deacon, T. W. 2011. *Incomplete Nature: How Mind Emerged from Matter*. WW Norton & Company.

Dennett, D. C. 1989. *The International Stance*. The MIT press.

Johnson-Laird, P. N. 1983. *Mental models: Towards a Cognitive Science of Language, Inference, and Consciousness*, volume 6. Harvard University Press.

Jonas, H. 1966. *The Phenomenon of Life: Toward a Philosophical Biology*. Northwestern University Press.

Jonas, H. 1968. Biological foundations of individuality. *International Philosophical Quarterly*, 8(2): 231–251.

Kauffman, S. A. 2002. *Investigations*. Oxford University Press.

Maes, P. 1993. Modeling adaptive autonomous agents. *Artificial Life*, 1(12): 135–162.

Maturana, H. R. and Varela, F. J. 1991. *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy of Science. Springer.

Moreno, A. and Etxeberria, A. 2005. Agency in natural and artificial systems. *Artificial Life*, 11(1-2): 161–175.

Murphy, J. P. and Rorty, R. 1990. *Pragmatism: From Peirce to Davidson*. Westview Press Boulder, CO.

Rohde, M. and Stewart, J. 2008. Ascriptional and 'genuine' autonomy. *BioSystems*, 91(2): 424–433.

Ryle, G. 1949. *The Concept of Mind*. University of Chicago Press.

Stewart, J., Gapenne, O., and Di Paolo, E. A. 2010. *Enaction: Toward a new Paradigm for Cognitive Science*. The MIT Press.

Thompson, E. 2004. Life and mind: From autopoiesis to neurophenomenology. a tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, 3(4): 381–398.

Varela, F., Thompson, E., and Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. MIT press.

Varela, F. J. 1979. *Principles of Biological Autonomy*. North Holland New York.

Weber, A. and Varela, F. J. 2002. Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2): 97–125.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0006

Enactivism and Freedom Education

Qing Li[*]
College of Education
Towson University
[*] *li[]towson.edu*

Ian Winchester
Faculty of Education
University of Calgary

Received November 2013; accepted September 2014; published Autumn 2014.

Abstract

In this paper, we argue, grounded on empirical evidence, that enactivism is a promising philosophical stance with great potential to address challenges brought by our rapidly changing world. We then propose Freedom Education, a new form of teaching and learning founded on the enactivist theory. After discussing what constitutes Freedom Education and what it is not, we recommend several principles to establish a learning world of freedom education.

Keywords: Enactivism; Education; Freedom Education; Cognition; Learninng.

Our world is changing

We live in a changing world. In his book titled *Exodus to the Virtual World*, Castronova (2007) describes how millions and millions of people are migrating to virtual worlds. People, for instance, immerse in a collective fantasy in massive multiuser online environments. In such a fantasy world, they may see or even build anything, whether a stately palace, a magnificent castle, or a peaceful landscape with ocean views. They may also see and interact, through typing, texting, talking, with other characters who can be either machine run (i.e. controlled by the system's artificial intelligence engines) non-player characters or avatars controlled by other human beings. They can do various trivial or odd things ranging from blacksmithing to practicing yoga skills in this virtual world just as if they were in a real world. Called "virtual worlding," such an increasingly popular practice signifies that these people have immigrated to the new land of the virtual worlds.

The number of people who have gone off to this virtual frontier is growing rapidly. For example, Second Life (SL) is a virtual world developed by Linden Lab and launched in 2003. The office SL website claims (Second Life 2013) that SL gives free membership and allows users, called residents, to interact, explore, meet others, socialize, participate in activities, create, do business, or travel throughout the world. Since its launching, over 36 million accounts were created with more than a million people visiting SL each month. In a similarly vein, according to the Wikipedia, over seven million subscribed to the World of Warcraft as of July, 2013. World of Warcraft has the highest number of subscribers of a Massively Multiplayer Online Role Playing Game (MMORPG) which gives it the Guinness World Record of the most popular MMORPG (World of Warcraft, n.d.). Minecraft is another virtual game world which many people inhabit. As of 2013, over 12,554, 000 people had bought the game (Minecraft 2013). While the numbers for these sorts of virtual worlds may rise and fall, new worlds are constantly emerging online.

At first glance, the spreading of the population in virtual worlds may seem trivial. However, as Castronova (2007) points out, the sheer quantity of people who spend so much time and energy immersed in the virtual worlds has significant impacts on every aspect of our life. For example, "Second Life (SL)" has its own economy and currency, the Linden Dollar, which can be exchanged with US Dollars. The Wikipedia (Economy of Second Life, n.d.) reports the following statistics: about sixty four thousand users made a profit in SL in February 2009. The SL economy grew 65% in 2009 to US $567 million, while the entire US virtual goods market is about $2.7 billion US dollars. Although no recent data on this aspect is available, the trend is obvious. Edward Castronova, in his book mentioned above, convincingly argues that our exodus to virtual worlds is forever changing our life both in virtual and real worlds. Such a change is so substantial that our current educational systems are facing serious challenges.

The Emergence of Participatory Culture

What is more, we are observing the emergence of a participatory culture. Technology advancement allows more opportunities for people to interact, collaborate, create, and share. To see following statistics about Facebook alone (Henrikson 2011): Facebook had 750 million users in 2011, which means one out of nine people in the whole world was using Facebook; The 2010 data show that every 20-minutes, people post close to 6 million wall posts, upload about 3 million photos, and write over 10 million messages on this social media site.

The significance and consequence of these developments are difficult to estimate at the current stage. Yet, most of us will agree that new technologies, especially the emergence of Web 2.0 tools, are changing our life fundamentally. Web 2.0 is a "perceived ongoing transition of the WWW from a collection of static websites to a full-fledged computing platform serving Web applications for end users" (O'Reilly 2005).

This trend of more and more people moving from digital consumers to creators means the increased civic engagement in culture, leading scholars (Jenkins et al. 2006) to believe that a participatory culture is emerging. Participatory culture is "a culture with relatively low barriers to artistic expression and civic engagement, strong support for creating and sharing one's creations, and some type of informal mentorship whereby what is known by the most experiences is passed along to novices. A participatory culture is also one in which members believe their contributions matter, and feel some degree of social connection with one another" (Jenkins et al. 2006: 3).

In a participatory culture, development of the 21st century skills, collaborative learning, creative expression of cultural diversity are placed at a premium (Jenkins et al. 2006). Access to this participatory culture, these authors argue, becomes a new hidden curriculum. Consequently, this brings significant challenges to education.

The Challenge of Participatory Culture and Virtual Worlds to Education

Whether it is the phenomena of exodus to the virtual world, or the shift towards the participatory culture era, the changes are so fundamental that it forces us to rethink education in principle and the current educational systems in general. As Dede (2008) claims, such fundamental shifts call for the reexamination of education because our traditional views about knowledge, expertise, and learning are being challenged.

Various approaches have been proposed to address such challenges: some scholars suggest improving the existing educational system by modifying current policy and pedagogy to help students develop the 21st century skills (Jenkins et al. 2006); others are more radical by suggesting that we should completely reconsider and redesign our formal educational systems (Dede 2008; Erneling 2010). Important questions such as: how we teach and learn to prepare students to become full participants in our contemporary world, how we are involved in a participatory culture, demand new educational theories and practices.

This paper, therefore, proposes Freedom Education, a new way of creating learning worlds grounded in enactivism. This paper is a further development of our earlier work published in the *British Journal of Educational Technology* (Li et al. 2010).

Enactivism and Freedom Education

Enactivism has recently emerged as a viable framework that provides a more encompassing philosophical stance accounting for learning and creation than other viewpoints (Li et al. 2010; Winn 2006). For example, constructivism, a philosophical viewpoint which has dominated the field of education in the last couple of decades, is grounded in two important perspectives. First, our personal world is one that is socially constructed by us collectively. Second, our personal world is individually constructed in that social context. Enactivism, however, suggests that beyond such constructed worlds, there are many situations in which the world is not constructed by us either collectively or individually, but rather formed out of an interaction between ourselves and our environment in such a way that both ourselves and our environment are transformed.

In this section, we discuss enactivism only briefly due to limited space available. The main idea of the following discussion about enactivism was first published in the *British Journal of Educational Technology* (Li et al. 2010). To provide readers with more fluid reading, we may or may not use quotation marks. Those who are interested can read our earlier paper (Li et al. 2010) for a detailed articulation of enactivism.

Enactivism has its roots in both phenomenology (Merleau-Ponty 1964a) and biological perspectives (Bateson 1972). On the one hand, enactivism is grounded in Merleau-Ponty's phenomenological view of ontological embodiment, which is based on the idea that "the world which is given in perception…is the concrete, intersubjectively constituted life-world of immediate experience" (Merleau-Ponty, 1964b: xvi). The enactivists' view of double-embodiment contends that

> the world is inseparable from the subject, but from a subject which is nothing but a project of the world, and the subject is inseparable from the world, but from a world which the subject itself projects. (Varela et al. 1991: 7)

Double-embodiment stresses that our body is both a lived structure to experiences and the setting for cognition (Varela et al. 1991). "Mindfulness meditation," a traditional Buddhist idea, has also influenced enactivists' thinking, in which our mind is placed in

> embodied everyday experience… [Our reflection] can change from an abstract, disembodied activity to an embodied (mindful), open-ended reflection. By embodied we mean reflection in which body and mind are brought together…[Refection] is not just on experience, but reflection is a form of experience itself and that reflective form of experience can be performed with mindfulness/awareness… (Varela et al. 1991: 27)

Embodiment is the "developing process" of our interaction with our world, including how we act, do, and perform in order to experience the environment in which we are living. In such a process, our thinking, behaviors, and sensors are all enmeshed and intertwined in helping us making sense of the experience (Johnson 1989).

On the other hand, the biological origin of enactivism includes systems theory and cognitive theory (Michie 2004). Enactivism describes living as systems that produce themselves endlessly (Reid 1995). Under this view, "living systems are not simply observation objects or interacting systems, but rather autonomous, self-contained, self-referencing and self-constructing closed systems" (Maturana & Varela 1980: v).

The idea that cognition is embodied has been discussed widely and tested with numerous examples and experiments. For instance, recent brain research indicates that the brain has a plasticity never dreamed of several decades ago (Diodge 2007). Before, brains were thought to have a fixed capacity and a limited number of cells and connections. What has been discovered, as explained thoroughly in Diodge's book (2007), is that people can recover completely from devastating strokes through an enactivist program that involves physical exercises as well as thinking about exercises. Such a program of both active physical movements and the mental processes of the movements enables these physical movements to be possible at a later time.

Results from research in both behavioral science and neuro-science demonstrate that some traditionally considered purely symbolic psychological phenomena in fact show perceptual effects (Black et al. 2012). For example, in the 1970s, Black and his team (Black et al. 1979) did an experiment examining the impact of perceptual effect on reading comprehension by asking people to read the following sentences:

1. *John was working in the front yard then he went inside.*
2. J*ohn was working in the front yard then he came inside.*

Evidently, sentence one is exactly the same as the sentence two except the word "went" is changed to "came". Yet, the participants took longer time to read the second sentence than the first one. The researchers later discovered that this difference of reading time was caused by the change of perspectives. Apparently, people develop a mental image when reading texts where they visualize in their head, the situation and the people being described. When the word "went" became "came," it changed the spatial perspective in the narrative. People therefore needed to adjust their point of view in their mental image, which resulted in longer reading time and more memory errors (Black et al. 2012). Examples like this demonstrate the inseparability of our body, mind, and the environment, the key idea that enactivism argues for.

Co-evolvement is an important concept of enactivism, which relates to the idea that cognition is a complex process of systems co-evolving with each other and the environments (Davis et al. 2000). For enactivists, cognitive system is the builder of meaning instead of only a vehicle for processing information (Baerveldt & Verheggen 1999).

The historical dualist debate considers knowledge either as presentations of reality or as individual agents learner developed inside their "inner-self." Enactivism challenges this debate by considering knowledge to be a domain of possibilities that emerges from "structured coupling" where systems are co-effecting each other in an ever-evolving world (Varela et al. 1991). We believe that knowledge "does not drive the actions of a living system but unfolds in events that evoke these particular actions (Fenwick 2000). Consequently, learners are believed to be an integral part of the context itself. Understanding, therefore, is embedded in action and based on both conscious and non-conscious knowing. Learning is not about gaining information, but an ongoing process of exploration about consciousness, self, context, and interactions of complex systems in order to adapt to the evolving environments" (Li et al. 2010).

The contemporary world in the ordinary sense is relatively stable. For example we have lived in a world with automobiles, telephones, airplanes and the radio for over a hundred years. While small changes have occurred in these things, this last century is arguably the most stable one we have ever had. In contrast to the relatively stable real world, the new virtual world is changing every second. We need not only to be able to keep up with it but also transform it into something better. What follows is the need for a different kind of education that begins with the notion that our world is not only changing rapidly but that we ourselves are too.

We, therefore, suggest that a new form that we term "freedom education," an educational approach grounded in enactivism, can provide solutions to the problems of education that the new virtual environments meld into our traditional ones. We argue that this approach enables us to address the problems we encounter in our current and presently transforming society. Before we delve into that, however, we first discuss the problems of nearly all contemporary learning theories and current educational systems.

WHAT IS THE PROBLEM?

In this section, we borrow Erneling's (2010) discussion to briefly describe the basic problems our current educational systems facing without repeatedly citing her work. We start the dialogue from the discussion of two basic ideas, which leads to the introduction of learning theories in computer related contexts. We then argue the need for a paradigm shift and propose "freedom education."

Creativity and Learning Theories

What is learning and what is the focus of learning theories? The Wikipedia defines learning as "acquiring new, or modifying and reinforcing, existing knowledge, behaviors, skills, values, or preferences" (Learning, n.d. para. 2). Most, if not all, contemporary learning theorists would agree that they mostly care about how learners change in their cognitive ability (i.e. can move beyond information provided) and such a change is not merely caused by chemical or physiological change, rote memorization, or mindless repetition. In addition, learning is not about creativity (Erneling 2010).

Let's compare two students Jason and John: Jason can remember the Pythagorean theorem and is able to regurgitate it. Yet, he flounders when asked to apply this theorem to solve any problems beyond the examples given. John, on the other hand, can solve new problems applying the theorem in new contexts. Therefore, John, but not Jason, has learned the Pythagorean Theorem because he can go beyond what is given—demonstrating a cognitive ability change.

Erneling (2010) asserts that while learning is all about using knowledge to solve new problems, creativity does not belong to the domain of learning or cognitive development, consequently cannot be explained by learning theories. Significant conceptual or other changes are the hallmark of creativity. "If we were in possession of a theory which could explain and predict radical change, the change predicted would already be present in the theory and just making it explicitly would not count as creativity" (Erneling 2010: 19).

With these two concepts suggested and how they appear to be in conflict, Erneling (2010) discussed how the ideas of productivity and educational framework pose a practical dilemma for educators. In this paper, we focus on the learning theories assumed by dominant pedagogical thinking on computers.

Learning Theories in Computer Related Contexts

According to Erneling (2010), a majority of research projects conducted in educational technology, in particular how computers can be used to promote learning seemingly accept the idea that infants learning to talk, walk, interact represents an exemplary learning situation: they are always interested, learning seems painless, and no explicit instruction is needed. This is not only reflected in various theories in the field of educational technology such as Papert's constructionism but more broadly in various learning theories. The view that learning is essentially fun and fun is more important than learning something new has been widespread. This might be termed *the standard picture of computer related pedagogy.*

She (Erneling 2010) argues that such infantilisation of education largely ignores the various different factors contribute to learning. Such differences include, for example, different cognitive skills and different social situations. As well, infantilised learning assumes that learning of everyday experience is the same as the learning of school subject matter knowledge, yet we know that everyday learning is 'natural' and the other is abstract, symbolic and conventional. Agreeing with Erneling's (2010) view about the problems of infantalization of education, we use her arguments to build our case. In the next 2 sections, we again borrow her ways of discussion focusing on the work of Piaget and Chomsky.

Piaget's Theory

Piaget's theory, especially his focus on qualitative development of human beings, has a fundamental impact on education. Most of our educational programs and instructional approaches are largely grounded in his idea that learning is most effective when children are developmentally ready. His theories propose that any cognitive change involves assimilation and accommodation. Such fundamental indiscriminative view therefore supports the approach that infant learning can be applied to all learning regardless of age, gender, or culture.

Grounded in the biological perspective, Piaget' theory assumes that all cognitive development are processes of biological adaptation to the environment (Piaget 1967). This fundamental assumption leads to the belief that all learning can mimic infant learning, consequently individualizing and infantilizing education. Piaget's famous cognitive development stage theory proposes that every child goes through four stages: sensory-motor, pre-operational, concrete operation, formal operational stages, from simple to complex and from concrete to abstract (Piaget 1967). At the first glance, this appeals contradictory to the infantilisation of education because it differentiates mental structures of different stages. Yet, diving deep down, as

Erneling (2010) argued, we see his two important points supporting infantilisation of education. First, Piaget's pervasive arguments suggest the intelligence of infants even if they cannot talk or walk. Secondly, his theory stresses that the same basic biological principles rule all cognitive development. That is, the development process follows the same assimilation and accommodation process. *Assimilation* refers to the process in which new materials are assimilated to our old structures, while *accommodating* describes that when new information cannot be fit into the old structure, new structure needs to be created to accommodate (Piaget 1967). In summary, the infantilisation of education is inspired by Piaget's theory, which claims that cognitive development processes remain the same regardless of contexts, that infants' learning provides the ideal model for any human learning.

Focusing on the Piaget-inspired view of learning, Erneling (2010) uses Seymour Papert's work of pro-technology research as well as anti-technology studies as examples to demonstrate that both the critics and the advocates of technology share similar assumptions: (1) all learning should be natural, which is equivalent to mastering biological and cognitive skills; (2) infants' learning presents an ideal learning situation. That is, both pro-technology and anti-technology researchers support natural learning. Erneling (2010) criticizes this idea of natural learning. She states that natural learning theories are essentially grounded in the learner's natural or biological, or innate, ability to learn. But what is involved in natural learning? For example, does natural ability develop following specific, law-like patterns with initial inborn instincts, as described by Piaget (Piaget 1980)? Or as Skinner (1974) suggested that the natural ability are open to change?

Another problem is that not everything new (i.e. going beyond experience) is learning. For example, not any random arrangement of variables is mathematics, not every utterance is language, unless they fit in to the norms of the specific cognitive activity. Learning is a normative practice within common frameworks. We are, therefore, limited on what we can go beyond existing experiences to new contexts and new directions. Natural learning theories ignore this issue, or at least do not deal with it explicitly. Instead, children are assumed to have an innate ability to judge and therefore can draw correct information from experience. Yet, we know that what is considered norm/knowledge in one culture may not be considered as norm or knowledge in another culture (Erneling 2010).

Chomsky's Innatism

Chomsky is another significant theorist, whose work has had significant impacts on learning theories. Since late 1950s, Chomsky's theories of innate language capacity and universal grammar largely contributed on research related to infant language learning. For Chomsky, who describes his picture as Cartesian, language learning requires very little exposure to language in order for a child to learn to speak and understand an infinite number of grammatical sentences (Chomsky 1966). This is not due to the child having learned an infinite number of sentences, but due to the fact that the child possesses a built in linguistic device that enables her to produce that infinite variety from a few examples of sentences in a language. The role of the environment is merely exposing individuals to limited human activities or providing background knowledge.

Chomsky disagrees strongly with Jean Piaget who does not believe the idea of innateness of mental structures (Piattelli-Palmarini 1994). Despite such disagreement, Erneling (2010) describes how these two share the same fundamental views about mental activity and cognition. For example, they hold the same assumptions and frameworks of mental activities, share the same philosophical view, dismiss empiricist developmental theories, and emphasize the learners' active role in cognitive development. Another significant similarity between Piaget and Chomsky is their shared belief about the biological foundation of all psychological activities and cognitive changes. They consider that all learning, thinking or perception, corresponds to some individual, private psychological activity. Cognitive growth is viewed mainly as an individual endeavor, minimizing the impact of the environment. They argue that what is true for learning in the infant and young child is true throughout the life of a human being. As far as we can see, Ernerling's strictures discussed above apply to both Piaget and Chomsky.

As demonstrated from the above discussion, Ernerling is one author who has seen the problems with contemporary learning theories. In particular she shows how they all essentially offer an account of how learning occurs understood as learning the normative standards and skills that are seen as contemporary ones. But none of these accounts offer anything useful with respect to creativity. Our employment of enactivism as the basis for our freedom education shows on the one hand why it is possible to nonetheless learn the normative standards and skills through the interactive picture offered by enactivism, but on the other hand suggests that at any point in that learning, creativity is necessarily a central part of the possibilities that are present in the context of that freedom.

FREEDOM EDUCATION

The difficulties Piaget's and Chomsky's works face in handing the fundamental philosophical problems (Erneling 2010) discussed by Erneling, are compounded by the challenges that education is facing with the confluence both of real and virtual worlds and of the shift to a participant culture mentioned at the outset above. Such difficulties call for new educational paradigms. We propose, therefore in this paper, Freedom Education grounded in enactivism, an emergent philosophical standpoint. Enactivism has been claimed by many, including the present authors, to provide a more embracing theoretical perspective that meets the current challenges (Davis et al. 2008; Li et al. 2010; Winn 2006).

Next, we sketch what freedom education would involve in a general sense, in an attempt to provide the first steps towards the freedom approach to education. While we employ several examples to instantiate our points, a more substantive account would involve moving beyond the scope of this paper.

Freedom education grounded in enactivism first and foremost stresses the significance of our world and our interaction with it. Although we acknowledge the importance of individuals' innate ability to cognition, our proposed Freedom Education differs significantly from both Piaget's viewpoints and Chomsky's standpoint. One of the biggest distinctions is that freedom education regards the environment as essential in cognition. Instead of thinking the environment only plays a subordinate or supplemental role, freedom education emphasizes that we cannot separate any human activity from the environment. Subsequently, our subjectivity, mind, and the environment are subsumed within larger systems rather than one dominating the other.

In freedom education, creativity is placed at the center of learning and related activities. From the Freedom Education's point of view, Chomsky's Cartesian picture (Chomsky 1966) is extended from language learning to learning in general. In this view, one characteristic of human learning is that from limited exposure to human activities of all kinds, most human beings are able to go on to solve an infinite number of similar problems without having to have been exposed to that problem situation and its solution before. This is the basis of human creativity, not only in language use but also in all human activities that have a rule-governed nature to any degree. One example of this is our ability to invent and play, in creative ways, an enormous variety of sports and games. Someone who has learned to play soccer can go beyond just following its rules, and in principle can make a creative advance in the sport spontaneously in an indefinite number of ways as well. This is true in very rule-bound activities like chess and most strikingly in a child's active playing with dolls or crayons.

The environment is crucial in freedom education such as an individual speaking a new language, or a young person mastering their first one. One only knows that the sentence one has just uttered is both grammatically and cognitively sound when those who have already learned to speak the language show that they understand. Beyond language acquisition, the experience of William Webb Ellis, a rugby player, provides another wonderful example. Webb Ellis was the inventor of rugby football who, when playing soccer with his classmates, picked up the ball and ran with it. What could happen at the time was that his classmates and teachers might have sent him to the principal for breaking the rules. But instead, they accepted what he had done as an innovation that would be useful for their game and adopted it. Such acceptance from the environment thus enabled the innovation (e.g. grab the soccer ball and run with it) to be normalized and therefore be recognized as legitimate knowledge creation. In that sense his innovation was grammatically and cognitively sound. Thus one might say that our enactivist "freedom education" paradigm essentially accepts that there might be biological or mental structures that lie behind human learning, but that these structures can only be exercised to effect if they are compatible with the environment in which they are operating. In a nutshell, Freedom Education accepts the view that cognition is biological and individual as Piaget claimed and at least partially innate (as Chomsky claimed only for language acquisition), but also powerfully social and cultural. The inseparability of mind, body, and the environment means that each aspect is equally important.

Because cognition is innate and biological, Freedom Education calls for an enabling world with a high degree of freedom for learner to explore, to investigate, to take risks, to innovate and to develop. Such a world also should contain built in rich stimuli to guide learners to the possible evolving patterns and to inspire creation. As well, the biological nature of cognition means that “doing” plays an essential role in learning. Physical and mental active enactment with the environment enables learners to uncover and interpret patterns and interactions in the process. There appear to be biological or mental structures that are built in to a human being that will not come in to play unless the environment is taken in to account. Thus freedom education has main requirements: the natural capacities and structures as well as the relation to the environment. Without both of these, learning cannot occur. What does this mean for a classroom? It means that we need to offer both the freedom of the learner to engage their natural powers or structures and such powers or structures must be active in an appropriate environment. This environment includes teachers, other learners and a general learning context.

Equally importantly, since cognition is also social and cultural, Freedom Education demands that such an enabling world contains carefully designed constraints. Such indwelling constraints can guide learners to coevolve with the environment towards a preplanned domain of possibilities. In addition, learners are not working in isolation, but rather interact with the world, the people in the world, and the content in the world. As suggested by Erneling (2010), learning and cognition are both afforded and constrained by three sets of conditions: (1) the biological body and brain of the agent, (2) the world where the agent situated, and (3) the content to be learned. The agent's physical body and brain have its special functionality and structure. The world is socially and culturally shaped with varied norms and principles, which is represented in activities like business and lawmaking. The content to be learned is represented in forms like manuscripts, records, movies or games, which are categorized as the "third world" by Popper (Popper 1978).

Freedom Education encourages free observation and free activity relating to tasks recognized by the learner as desirable to engage in or achieve. If a learner wants to be able to hit a top spin forehand in tennis, the learner will watch a variety of tennis players engaging in hitting top spins in the context of practice or a game until s/he feels s/he has a sense of what is involved. Then, s/he will try it her or himself without constraint or criticism. It is important to note that criticism does not mean the kinds of constructive criticism that provides feedback. We are not arguing against feedback, quite contrarily, we believe appropriate, minimal feedback helps learners enormously. Such feedback should be organically built in into the learning world with rich stimuli guiding learners' acts. Rather, we argue against the destructive criticism that distracts learners and limits learners' free exploration of the world around them. The kinds of minimal feedback needed, in this case, are often the fact that the ball does not go as the learner wanted it to, not the commentary from a coach.

Perhaps the greatest topspin in the history of tennis was that of Bjorn Borg, the five-time Wimbledon champion from Sweden. He practiced his topspin against his garage door for hours with nobody watching until he thought it was good enough. Then he played a game with an opponent and tried it out. Sometimes it worked and sometimes it did not. He went back to the garage door and worked on it again until he felt he had a better grasp of it. Borg's approach to mastering the topspin exemplifies an essential characteristic of "freedom education." In contrast, in a conventional tennis school, he would likely have been given instruction as to how to hold the racket, how to place his feet, how to swing the racket, and so on. As well, he would have practiced with an opponent from the beginning. One can learn topspin that way too, but it will never be a "Bjorn Borg" topspin.

Freedom Education is not new. It has been practiced in Buddhist education for centuries. In some traditions, the Buddhist initiates follow their guru step by step towards enlightenment. This following is not forced, but rather is freely chosen in the manner and at the time that the initiates wish. Enlightenment may or may not happen in the end. But when the initiates are done, they can in principle do everything precisely as the guru has demonstrated. This of course carries over into other educational areas where Buddhism in this tradition has had influence. For example in arts, such as playing a musical instrument like an violin or painting in a particular style, the guru works with the student until she or he has learned every location of the finger, every stroke of the bow until all the catalogue of pieces that the guru knows are now known also to the student. Nothing is ever forced. The student either follows or does not follow the guru’s lead and works on it on her or his own until s/he is satisfied.

The Buddhist tradition also includes an approach to enlightenment in which explicit teaching is not part of the activities between the guru and the students. The tradition of startling, often associated with both a Chinese and Japanese form of Zen Buddhism, is an example. In this practice, the Buddhist initiate gets an action or a puzzle or a koan. The initiate reacts to this action or puzzle or koan, which may lead to the enlightenment. To some degree, of course, this approach is closer to “freedom education” than the other approaches in that it encourages students to find their own way to the solution of the puzzles, often with the help of a guru with minimal feedback. Both this Buddhist tradition and our emphasis on the enactivist account of education form the foundation of our “freedom education.” Freedom Education begins with the notion that learners must find their own way to their learning, though a teacher is always a possible part of that way.

From our present point of view, the most famous example of freedom education in contemporary educational history is the Summerhill school A.S. Neill founded in England in the 1920s that is described in his *Summerhill* book and other writings. Till now, the Summerhill School is still open and has been well regarded. In 2007, the United Nations recognized the school for its excellence (Neill n.d.). Similarly, Bertrand Russell founded the Beacon School which has adopted a similar educational philosophy. In both these schools, children have been offered educational experiences similar to our envisioned "freedom education." That is, students have the freedom to choose what to learn and how to learn. However, constraints are also offered to make sure that the students have learned a few things important for their future lives. Students learn individually, yet work with one another through democratic activities in which the students discuss with one another and come to an agreement or, if necessary, vote. The schools have also adapted a democratic management approach where everyone has equal right to determine the rules (Lamb & Readhead 1992).

Not only can we find practices of Freedom Education in traditional Buddhist education in the Far East, or in modern society like the Summerhill School in England, it has also been undertaken recently in order to teach deprived children school subjects. The research project “Hole in the Wall” gives another example of what we refer to as *Freedom Education*. In 1999, Mitra, a professor in India, started to explore the potential of computers for children's learning (Mitra & Rana 2001). Mitra and his team installed an Internet connected PC (with a hidden camera) in a wall close to an urban slum New Delhi, and left. Nine months later, the team found amazing things: groups of kids were playing on the computer. Their data demonstrated that these kids, who had no formal schooling, self-taught each other how to surf the Internet and use the software by simply playing on the computer. Wondering whether this interesting result was only purely accidental, Mitra's team (Mitra et al. 2005) repeated the experiment in diverse locations. They installed computers with Internet connections in small rural villages, shantytown in urban cities, and remote poor areas in countryside towns. To their surprise, the results were incredibly consistent regardless of the geographic locations, contents to be learned or the student populations: the truth is, students can self-learn any subject by interacting with each other and with the computers. In fact, Mitra and his team discovered that ethnic minority children who had no prior biology background knowledge learned biology, children who did not know any English learned English, only through their self-monitored and self-regulated learning from a computer. Their experiments also included a comparison between regular school learning and this self-instruction approach, or Freedom Education. The results? The freedom education was as effective as any traditional formal classroom learning. More importantly, such freedom learning processes also helped improve students' social values and collaborative skills (Mitra & Dangwal 2010; Mitra et al. 2005).

These examples, from historic Buddhist education originating from the Far East, to the Summerhill school in England, to the "Hole in the Wall" project in India, may make you wonder what can happen when North American children learn school subjects from freedom education. If we rewind the time back to late 1980s and early 1990s, we can find another model of Freedom Education by looking at one elementary school, the Banded Peak Public School in the Rockyview school division in Canada, a school just outside of the city of Calgary, Alberta. The idea was to try to make modern digital computer technology and robotics ubiquitous in the school. The physical structure was architecturally designed with banks of computers in a circle out in the hallway. Whenever a child needed to access the computer, and that was up to the child, they would simply run out into the hallway and jump up to a chair with a computer in front of it. This is drastically different from traditionally approaches of how computers were integrated into the classroom. Typically, a computer or perhaps a few of them were located at the back of a classroom. Occasionally the teacher would assign a computer related task to the students and one at a time they would have to approach the computer, or the few, at the back of the room to engage in their teacher assigned task. Undoubtedly something was learned by this, but not very much. The unique approach in Banded Peak enabled the entire school to quickly become computer literate and the teachers who tolerated the degree of freedom which this presupposed were wildly rewarded with interested, knowledgeable and inquiring students.

All of these examples in this section suggest that the enactivist characterization of how we relate to our environment, and so learn with it, by it as well as for it is also a characterization of how we might best conceive our learning through freedom in our actions and our thoughts. The Buddhist seeking enlightenment, the child in A.S. Neill's school, the hole in the wall experiment in which uneducated children played with a computer and learned to use it, the breaking away from rigid rules and creating new games are all examples both of enactivism in action and of freedom education.

One may still wonder how Freedom Education relates to enactivism as compared to other philosophical stances such as constructivism. The most important thing, we argue, is that enactivist grounded Freedom Education assumes that the learner and her environment are in constant interaction, transforming one another. The learner is not merely "constructing" her world but is already embedded in a world that is changing in part because she herself is involved in the world and is herself changing. This is the essence of enactivism, but is also the precondition for freedom education in the sense that one is not constructing a world solely through intentional action but is acting freely and finds that oneself and one's world are themselves transforming/transformed and developing. This implies that one's body is in the world of one's doing and that one's mental life is also part of the world just as one's body is. The virtual world, as we emphasized in the beginning of this paper, interacts with us primarily through our senses and our minds, but is part of that environment that is not only partly shaping us but is also being shaped by us as we interact with it. Our relationship to that virtual world is not primarily through the many modes of the body as our everyday interaction with the world is, but it is just as intense and important.

There is, of course, something of a puzzle in the enactivist picture of how we relate to our environment as it tends to suggest that the normal mode is largely unstructured and unconstrained. We argue that, from the enactivist point of view, the normative world is part of that grand background or world in which a learner acts and lives. As all these examples demonstrated, the learner is chiefly transformed herself by “the largely tacit normative structure” without changing it very much at any one time and in the process picks up the essence of the normative world. It is definitely not a matter of “fixed knowledge, fixed approaches, fixed abilities and dead end education.” For the enactivist or for the supporter of freedom education, the possibility is always there that something radically new can occur in the relationship between the world and the learner much like in the rugby example. The emphasis therefore, is not on instruction, but rather on mutual coordination between our mind, body, and the world that provide the main tenets of freedom education.

What Sets Freedom Education Apart?

There are two things which we emphasize in Freedom Education that set it apart from other traditional educational beliefs. The first is that it is possible through such an approach to learn anything which our species is capable, culturally, of doing now. But second, and equally important, it promotes the creating and learning of things which nobody can yet do but which are within the realm of possibility. And, of course, we wish to encourage both of these human accomplishments as a standard result of education in general.

Creation and innovation: three aspects constitute the basic tenets of freedom education: (1) the learning of all the important things that our species has historically discovered, mastered and catalogued, (2) to offer constant attention to the possibility of the development, or the creation, of new ways of going about old things, and (3) the possibility of ways of going about completely new things. The first aspect is about learning existing knowledge, which is what our current educational systems are all about. The second and third aspects, however, set freedom education apart from other forms of education. Let us clarify the second and third aspects by discussing some examples from games. Lawn tennis began from a variation of "Real" or royal tennis in France, but as England had lots of lawns it was played out of doors on the grass. The game initially would have consisted of three basic strokes: a simple, flat forehand, a simple flat single-handed backhand and a simple underhanded serve. But someone, we do not know who, decided that an overhand serve was possible and tried it out. It wasn't against the rules and it clearly permitted an advantage to the server if skill could be gained at it, though it was more difficult than a simple underhanded serve. Thus a new development, and indeed a dominant one, occurred for the game of tennis. Today the best servers tend to win all of their own games and usually dominate the score.

We can also look at the creative developments in a team sport such a rugby football which is the precursor of Canadian and American football. Many believe that William Webb Ellis whom we mentioned before, born in 1823, with "fine disregard for the rules of football as played in his time took the ball in his arms and ran with it, thus originating the distinctive feature of the Rugby football game" (Willian Ellis n.d. para. 12). According to the Wikipedia, in 1870s, the unofficial story has it that a challenge game was played between two universities, McGill and Harvard, in Montreal and the Canadians were driven back to their own end. A Canadian player, following Webb Ellis's lead with a fine disregard for the rules, catching a lateral pass instead of kicking it, threw it down the field and the referee judged that the ball was in play where it landed. Subsequently when the ball was thrown down field players on the same team as the thrower would start out “on side” and try to catch the ball, thus originating the forward pass which dominates the Canadian and the American versions of what was relatively recently referred to as “rugby football” in North America today. The Americans embraced the “forward pass” right away but the Canadians still considered it illegal until 1929.

In both of these examples while the “new” action was not already recognized by the norm of the culture of the game, it was instantly recognizable as a possibility permitted by the context of the game and ultimately became a standard part of it. This sort of possibility is the central notion behind freedom education, in which we believe that the future we face as a species is such that we must constantly be connecting with our contexts and developing them as the needs of the future require an indefinite number of “Webb Ellis” like moves for our species to survive and prosper. Unlike some of the more recent thinkers, such as Erneling (2010), who put little emphasis on creativity, we consider it to be the central feature of the kind of education to which we aspire for all.

The story involving the Inuit of Labrador at Gander during the Second World War exemplifies the creativity of Freedom Education. This anecdote is from the father of one of us who served in air force intelligence and air traffic control at that time at Gander. During that time, Inuit of Labrador were brought to repair aircraft engines at the Gander airbase, the busiest airport on the planet at that time as it was involved in handling the protections of the convoys of boats to Britain from North America that had to pass through Nazi submarine patrols. The Inuit, who could not read English and had no training in aircraft engines, were able to repair sophisticated aircraft engines. Apparently, their own “Freedom Education” practices were crucial in their possessing this knack, which enabled them to creatively solve new problems.

In Freedom Education, innovation and creativity, instead of the traditional basic academic subjects, are placed at a premium to facilitate students’ emotional and physical well-being and intellectual abilities for independent judgment. Accordingly, self-actualization and self-understanding are two aspects being promoted. At the center of Freedom Education is the encouragement of exploration, puzzle-solving, as well as playful and spontaneous work—all of which are important for creation and self-renewal. Our central interest is on what is possible and potential, instead of making learners vulnerable prisoners to existing knowledge.

What Freedom Education Is Not

But what kinds of education do not count as “freedom” education? Essentially, almost all of our historical and traditional approaches to education, especially formal education, are outside the bounds of what we wish to argue for. For example, in the ordinary classroom or in university instruction worldwide we primarily depend on students listening to the teacher’s lecture and students, if they are able, taking notes. An exam will follow based on the lecture. To be sure we can learn a great deal this way. But what we learn is in the context of fixed knowledge, fixed approaches, fixed abilities and dead end education.

At the present time, smartphones, fantastically powerful handheld computers connected to the vast resources of the internet, are being confiscated in for fear that the children will text one another or check in with their friends on social networking sites. This is clearly just the wrong approach to handling the fact that now nearly everybody has their own personal computer in their hand. Practically any learning task can be augmented with the use of the available applications, of which there are now thousands and thousands. But for this we require an understanding of freedom education and what it can do. Robotics kids learn from playing Robotics freely.

How would an enactivist "freedom education" differ from traditional education? Imaging this extreme version of traditional education: a child sits at a desk with a book opened to a specific page in a room with blank walls. She follows the teacher's instruction on a typical task like add up the numbers on the page. The child has nobody to talk to other than the teacher. Adding more students to the room will turn this to a typical classroom. In contrast to such a traditional approach, in freedom education, emphasizing both the child’s powers and the environment, the child has an active teacher, active classmates both of whom she may communicate with, access to the internet with infinite possibilities and freedom to interact with them all as she sees fit. While there are some constraints in the sense that we expect there will still be a curriculum to be explored, the child's exploration of that curriculum will be indefinitely varied.

In this paper, we have borrowed some of Erneling’s ideas to establish our argument against current educational systems, consequently proposing freedom education based on enactivism. It is important, therefore, for us to clarify how freedom education differs from Erneling’s (2010) discursive education. Although we agree with much of Erneling’s criticism of Piaget’s theory and the infantilization of education, our proposed Freedom Education differs significantly from Erneling’s (2010) discursive education. For Erneling:

> All learning… is a discursive undertaking, cognitive change is always a social process in which both the form it takes and the content involved are culturally and historically varied. Learning and cognitive development involve the domestication, not infantilisation, of the learner. Natural enabling conditions

> are always involved, but the longer learning proceeds the less important these become…Acquiring a language is the most important precondition for taking part in symbolic activities, especially in schooling. But schooling also involves other skills, such as the ability to ascribe beliefs to others and meta-cognition, the second order ability to reflect on and criticize one's own and others' cognitive activities both publicly and privately…This involves the awareness of one's own beliefs in relation to the norms and standards set and agreed to by teachers with institutional authority. In this important sense, what is private and individual is secondary…School is in a sense the discursive activity par excellence. (172)

This long quotation paints the picture of discursive education as proposed by Erneling (2010). Although she has never explicitly described what discursive education is, her discussion apparently indicates that such discursive education offers a picture heavily weighted toward the teacher as the essential environment in the educational enterprise. Freedom education, we argue, is different from discursive education in two important ways. First, in freedom education, both teachers and important other speakers, as part of the environment, play a crucial role in learning. In addition, the freedom education world includes physical, social, cultural aspects of the environment. Therefore, what is individual, private, innate, social, cultural, public are all equally important. Unlike discursive education, which considers what is private and individual secondary, in freedom education, individual factors are placed at the same level of importance as social and cultural ones, with no hierarchy.

Secondly, discursive education considers the symbolic in the context of human discourse, mainly the teacher talking with the students, to be primary and more important the more sophisticated the student is as a speaker of their common language. Freedom education differs from discursive education in that teacher talk or using language to communicate is not the necessarily the only or the primary way for people to learn. For example, Inuit children learn to build kayak by watching their parents making a kayak from materials like sealskin, bones and sinew sewn. They may do this at various ages from say two to ten or fifteen. The children need not talk with the parents at all although they could ask about the process. But in the end the child can manage to build a kayak her or himself. The Hole in the Wall project provides another example for people to learn what Erneling (2010) refers to as "abstract" knowledge or school subjects like chemistry or technology. This project demonstrates that children can learn various subjects from playing and interacting with a computer without any help from any adult. They even learn a new language, e.g. English, from playing with a computer.

Finally, the child herself is engaged in creative activity with respect to a much wider world. This world includes not only the social and cultural environment the child situated, but also the physical and virtual world she resides. She enacts and interacts with such social, cultural world where physical and virtual spaces are intertwined, from which she learns.

Recommendations to establish a learning world of freedom education

We have discussed various examples of freedom education, from ancient Buddhist education, to Inuit people learning to build a kayak, Mitra's hole in the wall studies, to current students in Canada learning schools subjects. These examples demonstrate that freedom education leads to students' greater understanding. To summarize, we have the following recommendations for those who wish to transform a traditional classroom into one of freedom in the sense we mean.

1. The learning world of freedom education should contain rich stimuli with multiple sensory modalities that inspire students' curiosity to explore freely and intentionally. The learning process should also encourage and promote the use of bodily actions that are conceptually congruent with the knowledge being learned to enhance students' understanding (Black et al. 2012).

2. This learning world should also have built in opportunities, in enactivist terms "affordances" and "constraints," that are carefully crafted with the intention to enable learners to progress towards the possible learning goals in their own time and in their own way. The goals and the processes of learning themselves are not predetermined, but rather negotiated between the learner and her or his world along the way.

3. The inseparability among our mind, body, and the environment suggests the importance of bodily movement in cognition (Li et al. 2010). Consequently, the learning world can facilitate students' comprehension by encouraging students' direct experience of a phenomenon through activities like acting it out with their own body and then moving towards a more abstract understanding (Black et al. 2012). Similarly a child might act out with his or her own mind in the manner that many are engaging in during their recovery from strokes that the discovery of the astonishing plasticity of the brain.

4. Motivation and emotions should be taken into account when designing a learning world of freedom education. Motivation always has an emotional component, but motivation is of two basic kinds, external and intrinsic or internal. In Freedom Education we wish to emphasize intrinsic or internal motivation since the emotional component is always positive and strong. Thus we offer freedom of interaction with the environment for the learner who is both determining her or his goals and developing her or

his own approaches to the achievement of those goals in the context of the environment.

5. Creativity and innovative skills are placed at a premium when considering freedom education. Learning is definitely not a matter of fixed knowledge, fixed approaches, fixed abilities and at the fixed time. The main tenets of freedom education, rather, are about promoting and encouraging free explorations and innovative ways of learning so that one can best adjust to the world she or he is situated in.

Although we describe these general guidelines for creating a learning world of freedom education, we stress that freedom education grounded in enactivism does not prescribe particular forms of instruction. Freedom education is based on a systematic and ideationally driven approach for educators that uses the theoretical assumptions that underlie enactivism and is exemplified by a number of educational practices that have never been collected together before to understand human cognition and direct the establishment of the learning environment.

In conclusion, we believe that the enactivist proposition is a true autonomous theoretical proposition with a promising future. The above discussion with ample examples of freedom education grounded in enactivism demonstrates that enactivism can offer a coherent and holistic research framework for cognition and beyond.

References

Baerveldt, C., & Verheggen, T. 1999. Enactivism and the experiential reality of culture: Rethinking the epistemological basis of cultural psychology. *Culture & Psychology 5*(2): 183-206.

Bateson, G. 1972. *Steps to an ecology of mind*. NY: Ballentine Books.

Black, J., Segal, A., Vitale, J., & Fadjo, C. 2012. Embodied cognition and learning environment design. D. Jonassen & S. Land, eds. *Theoretical Foundations of Learning Environments* (2nd ed.: 198-223. New York: Routledge: Taylor & Francis Group.

Black, J., Turner, T., J., & Bower, G. H. 1979. Point of view in narrative comprehension, memory and production. *Journal of Verbal Learning and Verbal Behavior, 18*: 187-198.

Capra, F. 1996. *The web of life: A new synthesis of mind and matter*. London: Harper Collins.

Capra, F. 2002. *The hidden connections: A science for sustainable living*. London: Haper Collins.

Castronova, E. 2007. *Exodus to the Virtual World: How Online Fun is Cahning Reality*. New York: Palgrave MacMilland.

Chomsky, N. 1966. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper & Row.

Davis, B., Sumara, D., & Luce-Kapler, R. 2000. *Engaging minds: Learning and teaching in a complex world*. Mahwah, NJ: Lawren Erlbaum Associates.

Davis, B., Sumara, D., & Luce-Kapler, R. 2008. *Engaging minds: Changing teaching in complex times.* 2nd ed.. Mahwah, NJ: Lawren Erlbaum Associates.

Dede, C. 2008. A Seismic Shift in Epistemology. *EDUCAUSE Review, 43*(3), 80-81.

Diodge, N. 2007. *The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science*. NY: Penguin Group.

Economy of Second Life. n.d.. Wikipedia Retrieved Oct. 30, 2013, from

http://en.wikipedia.org/wiki/Economy_of_Second_Life

Erneling, C. 2010. *Towards discursive education: Philosophy, technology and modern education*. Cambridge: Cambridge University Press.

Fenwick, T. 2000. Expanding conceptions of experiential learning: A review of the five contemporary perspectives on cognition. *Adult Education Quarterly, 50*(4): 243-273.

Henrikson, J. U. 2011. The growth of social media: An infographic. *Search Engine Journal*. Retrieved from http://www.searchenginejournal.com/the-growth-of-social-media-an-infographic/32788/

Jenkins, H., Clinton, K., Purushotma, R., Robison, A., & Weigel, M. 2006. Confronting the challenges of participatory culture: Media education for the 21st century: 66. Washington, DC: The John D. and cathering MacArthur Foundation.

Johnson, M. 1989. Personal practical knowledge series: Embodied knowledge. *curriculum Inquiry, 19*(4): 361-377.

Lamb, A., & Readhead, Z. 1992. *The New Summerhill - A. S. Neill*. London: Penguin Books.

Learning. n.d.. Wikipedia. Retrieved from http://en.wikipedia.org/wiki/Learning

Li, Q., Clark, B., & Winchester, I. 2010. Instructional Design and Technology Grounded in Enactivism: A Paradigm Shift? *British Journal of Educational Technology, 41*(3): 403-419.

Maturana, H. R., & Varela, F. J. 1980. *Autopoiesis and cognition: The realization of the living*. Dordrecht, Holland: D. Reidel Publishing Company.

Maturana, H. R., & Varela, F. J. 1987. *The tree of knowledge: The biological roots of human understanding*. Boston: Shambhala.

Merleau-Ponty, M. 1964a. *The primacy of perception and other essays*. Evanston, IL: Northwestern Universit Press.

Merleau-Ponty, M. 1964b. *The primacy of perception and other essays.* Evanston, IL: Northwestern Universit Press.

Michie, M. 2004, April. *Teaching sciece to Indigenous students: Teachers as culture broker or is it something else?* Paper presented at the National Association for Research in Science Teaching, Vancouver, Canada.

Minecraft (Producer. 2013. Minecraft Statistics. Retrieved from

https://minecraft.net/stats

Mitra, S., & Dangwal, R. 2010. Limits to self-organising systems of learning - The Kalikuppam experiment. *British Journal of Educational Technology, 41*(5): 672-688.

Mitra, S., Dangwal, R., Chatterjee, S., Swati Jha, Bisht, R. S., & Kapur, P. 2005. Acquisition of computing literacy on shared public computers: Children and the 'Hole in the Wall' *Australasian Journal of Educational Technology, 21*(3): 407-426.

Mitra, S., & Rana, V. 2001. Children and the Internet: Experiments with minimally invasive education in India. *British Journal of Educational Technology, 32*(2): 221-232.

Neill, A. S. n.d.. *wikipedia*. Retrieved from http://en.wikipedia.org/wiki/A._S._Neill

Piaget, J. 1967. *Biology and knowledge*. Edinburge: Edinburgh University Press.

Piaget, J. 1980. The psychogenesis of knowledge and its epistemological significance. P. Piattelli-Palmarini, ed. *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*. Cambridge, MA: Harvard University Press.

Piattelli-Palmarini, M. 1994. Ever since language and learning: Afterthoughts on the Piaget-Chomsky debate. *Cognition & emotion, 50*: 315-346.

Plato. 1981. *Five Dialogues* (G.M.A. Grube, Trans.. Indianapolis: Hackett Publishing Co. Inc.

Popper, K. 1978. *Three Worlds*. Paper presented at the The Tanner Lecture on Human Values, Michigan: University of Michigan.

http://tannerlectures.utah.edu/lectures/documents/popper80.pdf

Reid, D. 1995. Enactivism Retrieved Nov. 20, 2006, from

http://plato.acadiau.ca/courses/educ/reid/enactivism/EnactivismDef.html

Second Life. 2013. *Wikipedia*. Retrieved from http://en.wikipedia.org/wiki/Second_Life - cite_note-nwn.blogs.com.2Fnwn.2F2009.2F03.2Fmillion.html-24

Skinner, B. F. 1974. *About bahaviorism*. NY: Random House.

Varela, Thompson, E., & Rosch, E. 1991. *The embodied mind: Cognitive science and human experience*. Cambridge, MA.: Massachusetts Institute of Technology Press.

Willian Ellis. n.d.. Wikipedia. Retrieved from

http://en.wikipedia.org/wiki/William_Webb_Ellis

Winn, W. 2006. Functional contextualism in context: A reply to Fox. *Educational Technology Research & Development, 54*(1): 55-59.

World of Warcraft (Producer. n.d., Oct. 27, 2013. Wikipedia Retrieved from http://en.wikipedia.org/wiki/World_of_Warcraft

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0007

The coherence of enactivism and mathematics education research: A case study

David A Reid
University of Bremen
dreid[]math.uni-bremen.de
Received December 2013; accepted August 2014; published Autumn 2014.

Abstract

This article addresses the question of the coherence of enactivism as a research perspective by making a case study of enactivism in mathematics education research. Main theoretical directions in mathematics education are reviewed and the history of adoption of concepts from enactivism is described. It is concluded that enactivism offers a 'grand theory' that can be brought to bear on most of the phenomena of interest in mathematics education research, and so it provides a sufficient theoretical framework. It has particular strength in describing interactions between cognitive systems, including human beings, human conversations and larger human social systems. Some apparent incoherencies of enactivism in mathematics education and in other fields come from the adoption of parts of enactivism that are then grafted onto incompatible theories. However, another significant source of incoherence is the inadequacy of Maturana's definition of a social system and the lack of a generally agreed upon alternative.

Keywords: enactivism; biology of cognition; mathematics education; theories of learning; autopoiesis; cognitive systems; social systems.

Introduction

Is enactivism a coherent and promising research perspective? Is it a conception, or maybe a framework for research? What is its future? In this article I will address these questions by making a case study of enactivism in my field, which is mathematics education research. I will review the main theoretical directions in mathematics education and the history of adoption of concepts from enactivism. I will consider whether enactivism provides a sufficient theoretical framework for research in mathematics education and whether enactivism is coherent in general.

What is Enactivism?

Enactivism exists in the eyes of observers, and so any discussion of it must begin with a description of the authors' observations. For me enactivism is fundamentally a theory about autopoietic systems and the biology of cognition. My first encounter with it was *Tree of Knowledge* (Maturana & Varela 1987) and I read backwards and forwards from there. Maturana's writing has been more significant to me than Varela's, especially 'Everything said is said by an observer' (Maturana 1987) and 'Reality: The search for objectivity or the quest for a compelling argument' (Maturana 1988). Hence, when I write about enactivism I am thinking of the ideas of Maturana and Varela up to and including *Tree of Knowledge,* plus material from *The Embodied Mind* (Varela, Thompson & Rosch 1991) and later writing as far as it overlaps with Varela's prior work with Maturana. There are clearly many connections between enactivism and the work of Bateson, McCullogh, von Foerster and Lakoff, for example, but I do not see their work as defining enactivism itself.

I realise this reading of enactivism will only overlap and not coincide with that of many readers. After all, the words 'enaction' and 'enactive' were used first by Varela, Thompson and Rosch (1991) and Maturana never uses them, so my emphasis on Maturana's ideas may seem odd. And I am also paying little attention to Varela's later elaborations of enaction and the connections he makes to the phenomenology of Merleau-Ponty and Buddhist mindfulness/awareness practice. On the other hand, the description of ena-ction in *The Embodied Mind* depends on a number of concepts, such as *embodiment*, *perceptually guided action*, *recurrent sensorimotor patterns, structure determinism* and *operational closure,* that were developed during Varela's work with Maturana in the early 1970s, and first described in many joint and individual publications, most notably *Autopoiesis and Cognition* (Maturana and Varela 1980a) and *Tree of Knowledge*. And most importantly, the key isight of enactivism, it seems to me, is the founding of cognition in biology, and that insight is best reflected in the collaborative work of Maturana and Varela and Maturana's work since then.

Theoretical frameworks in mathematics education

Enactivism was introduced into mathematics education at a time when the main theoretical debate concerned how to describe the social interactions between individuals.

> Two general theoretical positions on the relationship between social processes and psychological development can be identified in the current literature. These positions frequently appear to be in direct opposition in that one gives priority

> to social and cultural processes, and the other to the individual autonomous learner. (Cobb & Bauersfeld 1995: 3)

Psychological theories, giving priority to the individual autonomous learner, entered mathematics education at its very beginning as an academic field. Sociocultural theories came much later, and their introduction created significant tension.

Mathematicians had taught mathematics and written mathematics textbooks for millennia before mathematics education emerged as an academic discipline. A possible marker of its origin is the first issue of the journal *L'Enseignement Mathématique* in 1899, which included both an article on mathematics teaching by the famous mathematician Henri Poincaré, as well an article on "scientific pedagogy" by Alfred Binet, the director of the Laboratory of Experimental Psychology at the Sorbonne (Kilpatrick 1992: 7). This bringing together of mathematics and psychology established the theoretical basis for the field of mathematics education.

After its birth as an academic field at the beginning of the twentieth century, the further development of mathematics education was disrupted by the interruption of international collaborations caused by the two world wars. After the post war recovery, the field was revitalised by the introduction of the genetic epistemology of the psychologist Jean Piaget (1896-1980, a student of Binet). This became the dominant theoretical framework for the rest of the century. The founding fathers of mathematics education from this time, Zoltan Dienes (1916-2014), Ephraim Fischbein (1920-1998), Caleb Gattegno (1911-1988) and Richard Skemp (1919-1995) all trained as both mathematics and psychologists, except Gattegno who collaborated directly with Piaget. The importance of psychology in mathematics education is also marked by the founding in 1976 of the International Group for the Psychology of Mathematics Education, whose annual conference became the main research conference in the field.

Even at this time, however, there was a recognition that mathematics education must also consider the contexts in which learning takes place. Heinrich Bauersfeld recalls:

> From the very beginning [of PME] I was unhappy with the exclusive concentration on Psychology only, which meant focusing on the individual and neglecting the social dimensions of the complex teaching-learning processes. Research on the complex problems of learning/teaching-processes and of teaching teachers to teach mathematics will not arrive at helpful constructive information as long as such vast domains as language, human interaction (not the usual psychological interaction of variables!) and rich case studies are neglected and/or treated by inadequate research methods. (Nicol et al. 2008).

By the 1980s "Researchers were taking the social and cultural dimensions of mathematics education more seriously (Kilpatrick 1992: 30). This meant that the theories of psychology had to be supplemented both by teachers' experiences and by theories from other disciplines. "Practitioners were increasingly becoming members of the interdisciplinary groups needed to help research link the complexity of practice to theoretical constructs. The techniques and concepts used by anthropologists, sociologists, linguists and philosophers proved helpful in that task" (Kilpatrick 1992: 31).

However, as I will discuss below, sociocultural theories did not fit easily into a field dominated by psychological perspectives, and as indicated in the Cobb and Bauersfeld quotation above, the two approaches often seemed to be opposed. Researchers in both camps critiqued the perceived failings of theories used by the others. This was the context in which concepts from enactivism first entered mathematics education.

Enactivist ideas in mathematics education

In mathematics education the growing influence of enactivist ideas can be traced historically through references to the work of Maturana and Varela. Enactivist ideas have been introduced into mathematics education four times. The first references come from radical constructivists who sought to incorporate the concept of *consensual domains* in order to address criticisms that radical constructivism did not address learning in social situations. Tom Kieren's work then introduced the full range of enactivist concepts, and he was the first to use the work 'enactivist' to describe his research. At about the same time, the concept of embodied cognition began to be used by a number of other researchers with interests in bodily metaphors and gestures in mathematics. Finally, the concept of autopoiesis, as reframed by Niklas Luhmann, has been used by Heinz Steinbring to describe interactions in mathematics classrooms.

Radical constructivism and consensual domains

The first influence of enactivism (prior to the coining of the term) came from an attempt to integrate social elements into the radical constructivism of Ernst von Glasersfeld. In the late 1980s, Paul Cobb (a graduate student of von Glasersfeld's collaborator Les Steffe) began to refer to Maturana's concept of *consensual domain* (1980b, 1978a) in order to describe a world view or belief system shared by many individuals (see e.g., Cobb 1985, 1986). Von Glasersfeld (1989) himself also refers to Maturana (1980a) to account for the possibility of communication between individuals, which arises "in the course of protracted interaction, through mutual orientation and adaptation" (Glasersfeld 1989: 132). This led to Maturana's work being seen as a part or a type of radical constructivism. For example, Konold & Johnson refer to "the radical constructivism of Maturana and von Glasersfeld" (1991: 10) and Krainer includes Maturana on a list of those putting forth "constructivist positions" along with Piaget and von Glasersfeld (1993: 69).

Critics often claimed that radical constructivism was unable to explain social phenomena (Ernest 2010: 41), and Maturana's concepts of consensual domain, domain of interpretation, isomorphic structures and mutual adaptation continue to be used by radical constructivists to address this criticism (e.g.,Thompson 2008). However, though von Glasersfeld and Maturana were friends, they disagreed on some fundamental points, such as structure determinism and the system's environment (Glasersfeld 1991, Kenny 2007). The radical constructivists, thought they were the first to apply Maturana's ideas in mathematics education, cannot be said to be using enactivism, as they pick and choose concepts, seeking those that give radical constructivism a language to describe social phenomena.

Tom Kieren

Tom Kieren has been described as "one of a very few pioneers of enactivism within the mathematics education community" (Kieren & Simmt 2009: 28). In the mid 1980s he picked up *Autopoiesis and Cognition* (Maturana & Varela 1980) and was strongly influenced by what he read. He also encountered enactivist concepts through a paper Maturana delivered with Karl Tomm at a conference of the Department of Psychiatry at the University of Calgary, on "Languaging and the emotion flow" in 1986 and through *Tree of Knowledge* (Maturana & Varela 1987). Kieren was already well acquainted with radical constructivism, but found something new and exciting in Maturana's ideas. In 1988 he discussed this work with his colleague Susan Pirie and this led to the first publication in mathematics education that makes extensive use of Maturana's ideas, Pirie and Kieren (1989). They use the concepts of recursion in knowing, knowing as effective action as seen by an observer, autopoiesis, consensual coordination of action, and the aphorism "everything said is said by an observer" as the basis for a theory of mathematical understanding that has come to be known as the Pirie-Kieren model.

In 1994 Kieren published a reaction to two papers by radical constructivists in the *Journal of Research in Childhood Education.* In it he uses the word "enactivist" for the first time in the mathematics education literature; the word was coined in *The Embodied Mind* (Varela, Thompson & Rosch 1991). Kieren (1994) describes enactivism as a position on cognition that includes the concepts of structure determinism, structural coupling, bringing forth a world, observer dependence, satisficing, and co-emergence.

In the 1990s Kieren supervised a number of graduate students who went on to use enactivist ideas in their work. These include Judy Barnes, Brent Davis, Lynn Gordon Calvert, Elaine Simmt, and myself. Almost all include in their doctoral dissertations extensive descriptions of enactivist ideas (Barnes 1994; Davis 1994; Gordon Calvert 1999; Reid 1995; Simmt 2000).

Of these graduate students, Davis has published the most extensively and has done much to spread enactivist ideas to the mathematics education community. His doctoral dissertation (1994) was the basis for his first book (1996), in which enactivist ideas are presented in detail. He places special emphasis on enactivism's denial of the mind/body split, and the entity/environment split (1996: 77). He also refers to autopoiesis, co-emergence, embodied cognition, double embodiment, knowing as being and doing, and structure-determinism. Davis also published several related journal articles (Davis 1995, 1997; Davis, Sumara & Kieren 1996) in which he provides descriptions of enactivism as an interpretive framework. For mathematics education researchers such as Begg (1999, 2002, 2013) and Samson and Schafer (2010, 2011, 2012), Davis is a primary reference for enactivist theory. However, Davis himself later came to regard enactivism as a variety of complexity science, and changed his focus to applications of complexity science in general to educational research (see, e.g., Davis 2004, Davis & Sumara 2006).

Three other doctoral students, Elaine Simmt, Lynn Gordon Calvert, and myself, along with Kieren constituted the Enactivist Research Group at the University of Alberta. Together we presented enactivism informed interpretations of shared data from four perspectives (Gordon Calvert, Kieren, Reid & Simmt 1995; Kieren, Gordon Calvert, Reid & Simmt 1995; Kieren, Simmt, Gordon Calvert, & Reid 1996). Key concepts used in those presentations were coemergence, structural determinism, autopoiesis, and double embodiment, and enactivism provided not only the interpretative frame but also the methodology.

Both Simmt and Gordon Calvert became professors at the University of Alberta, and doctoral work with an enactivist framework continues to be done there. The research of Joyce Mgombelo, Immaculate Namukasa, Jerome Proulx and Martina Metz at the University of Alberta has continued the enactivist tradition begun by Tom Kieren.

The Bristol School

After completing my dissertation (Reid 1995) I continued to make use of a wide range of enactivist ideas, especially Maturana's (1988) concept of an emotional orientation, in my research on the development of students' reasoning towards mathematical proof (see, e.g., Blackmore, Cluett, & Reid 1996, Reid 1996, 1999, 2002ab; Reid & Drodge 2000). Enactivism continues to provide the methodological underpinnings of my research into teachers' pedagogies (see http://www.acadiau.ca/~dreid/OT/).

In 1995 I met Laurinda Brown in a PME discussion group organised by Rafael Núñez and Laurie Edwards (see below). Brown had been using the work of Gregory Bateson (1972, 1979) extensively but had not discovered Maturana and Varela. I had not yet read Bateson's work. After exchanging reading lists, we began a long collaboration. With our colleagues Vicki Zack and Alf Coles we made use of enactivism as a theoretical framework, and also incorporated Bateson's work and later the work of Antonio Damasio (1994). His somatic marker hypothesis provided us with a neurological explanation for phenomena described by Maturana and Bateson (see, e.g., Brown & Reid 2002, 2003, 2004, 2006; Brown, Reid & Coles 2003; Reid & Brown 1999; Reid, Brown, & Coles 2001).

Brown introduced a number of graduate students to enactivist ideas, and two in particular picked them up and used them extensively. Maria Lozano completed her dissertation in 2004. She examined algebraic learning in the transition from arithmetic to algebra, using enactivism as both her methodology and theoretical framework (Lozano 2004). Alf Coles collaborated with Brown over many years, on research and publications (e.g., Brown and Coles 1997, 2000, 2008, 2010) that "adopted an enactivist epistemology and methodology" (Coles 2011: 18). In his doctoral dissertation (Coles 2011) he examined the patterns of communication in classrooms and teacher meetings from an explicitly enactivist perspective.

Embodied mathematics

The year 1995 can be seen as the birth year of another important theoretical framework in mathematics education, closely related but not identical to enactivism. In that year Stephen Campbell and A. J. Dawson published a paper on 'Learning as Embodied Action' (Campbell & Dawson 1995) which draws strongly on *The Embodied Mind* (Varela, Thompson & Rosch 1991). Rafael Núñez and Laurie Edwards in the same year presented a paper (Edwards & Núñez 1995) and organised a discussion group at PME on non-objectivist cognitive science (Núñez & Edwards 1995), drawing heavily on *The Embodied Mind* as well as the work of Lakoff (1987) and Johnson (1987) on body based metaphors in cognition. Since then embodied mathematics has emerged as a significant theoretical frame in mathematics education. It posits that all human cognition is embodied, that is "every subjective sensation, memory, thought, and emotion—anything at all that any human being can ever experience—is in principle enacted in some objective, observable, way as embodied behaviour." (Campbell 2010: 313). Three threads can be discerned within the theoretical framework of embodied mathematics.

One thread, which Campbell (2010) calls ‘mathematics educational neuroscience’, seeks to investigate mathematics learning using neuroscientific tools such as eye-tracking and brain scans. In this thread the ideas of Varela serve chiefly as a starting point to justify examining cognition in terms of observable bodily changes.

A second thread builds chiefly on the work of Lakoff. A basic text is *Where mathematics comes from: How the embodied mind brings mathematics into being* (Lakoff & Núñez 2000) in which they argue that mathematical concepts, even quite abstract concepts, are always based on bodily experiences, through metaphors. Researchers (e.g., Ferrara 2003; Oehrtman 2003) study the metaphors involved in students’ understandings of mathematical concepts such as functions, limits, and sets. No reference is made to the work of Varela, except indirectly through references to publications of Edwards and Núñez.

A third thread focusses on the use of gestures in mathematics education. This thread can be represented by a special issue of the journal *Educational Studies in Mathematics*, (Radford, Edwards & Arzarello 2009) in which embodied mathematics is used in combination with semiotics to research the role of gestures in mathematical thinking and communication. In this work Varela’s ideas play a limited role, acting mainly as a reference for the concept of embodied cognition.

As noted above, the theoretical framework of embodied mathematics is related to enactivism, but distinct from it. *The Embodied Mind* (Varela, Thompson & Rosch 1991) is a key reference for this school of research, but other work by Varela is rarely cited, and work by Maturana is almost never cited by researchers in the area.

Autopoietic social systems

The enactivist idea of autopoiesis found its way into mathematics education through one other channel, the sociological work of Nicholas Luhmann, which was applied to mathematics education by Heinz Steinbring. Luhmann (1986, 1995, 1997) considers social communication to be an example of an autopoietic system. In order to do so he generalises Maturana’s concept of autopoiesis to apply to non-living systems that also have the properties of being self-organising and operationally closed. Within this broader conception of autopoietic systems he identifies three types: living systems, psychic systems and social systems. The elements of psychic and social systems are not physical but based on meaning, in consciousness and communication respectively.

The relationship between psychic conscious mental and social communication systems is an important theme in Luhmann's work, and in its application to mathematics education.

> Communication systems and mental systems (or consciousness) form two clearly separated autopoietic domains.... These two kinds of systems are, however, closely connected to each other in a particular tight relation and mutually form a 'portion of a necessary environment': Without the participation of consciousness systems there is no communication, and without the participation in communication, there is no development of the consciousness. (Baraldi et al. 1997: 86, translated by Steinbring 2005:320)

Steinbring uses Luhmann's concept autopoietic social systems composed of communications to analyse episodes in mathematics teaching (see, e.g., 1999, 2005). This approach has had only a limited influence in mathematics education.

Summing up

Enactivism has become recognised as a theoretical framework used in mathematics education, and it is interesting to see how it has been summarised by overviews of theories in the field.

Mason and Johnston-Wilder include enactivism in the *Fundamental Constructs in Mathematics Education* (2004). Key concepts cited are 'action is knowledge and knowledge is action" and 'everything said is said by an observer.' The enactivist idea of 'bringing forth a world' is seen as "entirely compatible with von Glasersfeld's *radical constructivism*" (Mason & Johnston-Wilder 2004: 71), and radical constructivists such as Cobb, Yackel and Wood are described as having "taken up the enactivist view" (Mason & Johnston-Wilder 2004: 72). The work of Campbell and Dawson (1995) is cited a particular example of the enactivist approach, with the idea of stressing and ignoring being presented as central. Much of Mason and Johnston-Wilder's summary seems to have been based on Campbell and Dawson's article, which appeared in a book edited by Mason. The body of work inspired by Tom Kieren is represented by a brief quotation from Davis, Sumara and Kieren (1996) which is seen as exemplary of "radical enactivism" (Mason & Johnston-Wilder 2004: 73), in which learning and action are identified.

In the recent collection *Theories of Mathematics Education* (Sriraman & English 2010) Paul Ernest contrasts four "philosophies of learning": 'simple' constructivism, radical constructivism, enactivism and social constructivism. Ernest (2010) cites *The Embodied Mind* (Varela, Thompson & Rosch 1991) as the original influence that brought enactivism into mathematics education and he lists autopoiesis and cognition-as-enaction as key concepts. He cites Lakoff and Johnson's work, as transmitted by Lakoff and Núñez, as a "source of enactivism" (Ernest 2010: 42). Ernest sees enactivism as "not so very different from Piaget's epistemology and learning theory and the radical constructivism to which it gave birth" (2010: 42). What Ernest sees are distinct in enactivism is the role of metaphor contributed by Lakoff. This suggests that it is the embodied mathematics research that has been Ernest's main source of information about enactivism. Another clue to this is his comment "What both enactivism and radical constructivism appear to share is the subordination of the social or the interpersonal dimension" (2010: 43).

Mathematics: Mind or Society?

Recall that when enactivist ideas were first introduced into mathematics education a fundamental tension existed between psychological theories or sociocultural theories. This tension was addressed in mathematics education in a number of ways.

As noted above, radical constructivists borrowed Maturana's concept of consensual domain in order to be able to address social aspects of learning. The concept of consensual domain was only referred to, however, and the full implications of incorporating the concept into radical constructivism were never, as far as I know, worked out. Given the fundamental ontological differences between radical constructivism and enactivism, it may not be possible to truly integrate the concept of consensual domain into radical constructivism. Adopting the concept of consensual domains without adopting the ontological basis for their existence would give rise to an incoherence in radical constructivism. This can only be resolved by either adopting Maturana's position on reality (effectively changing radical constructivism into enactivism) or by dropping the concept of consensual domains, leaving radical constructivism again open to the critique that it does not address learning in social contexts.

Cobb and Bauersfeld (1995) take a different approach, of employing both a radical constructivist framework as well the interactionist perspective developed by Bauersfeld (1980).

> We arrived at the conclusion that psychological and sociological perspectives each tell half of a good story. What was needed was a combined approach that takes individual students' mathematical interpretations seriously while simultaneously seeing their activity as necessarily socially situated. (Cobb & Bauersfeld 1995 p. ix)

They "seek to transcend the apparent opposition between collectivism and individualism by coordinating sociological analyses of the microculture established by the classroom community with cognitive analyses of individual students' constructive activities" (Cobb & Bauersfeld 1995: 7). However, they are explicit that "this coordination does not … produce a seamless theoretical framework" (Cobb & Bauersfeld 1995: 8). It is a coordination of approaches, not a single combined approach.

> Instead, the resulting orientation is analogous to Heisenberg's uncertainty principle. When the focus is on the individual, the social fades into the background, and vice versa. Further the emphasis given to one perspective or the other depends on the issues and purposes at hand. Thus … there is no simple unification of the perspectives. (Cobb & Bauersfeld 1995: 8)

Another approach begins with a sociocultural perspective and attempts to integrate psychological elements. The main example of this approach is social constructivism, which built on the work of the Belorussian psychologist Vygotsky. Partly out of a genuine interest in how thinking is related to language and society, and partly because the political and intellectual context of the Soviet Union in the 1930s demanded a Marxist element in any theory, Vygotsky developed a theory that pays explicit attention to how social and cultural patterns of interaction shape thinking. Beginning in the 1970s Vygotsky's ideas were adopted in mathematics education as an alternative, called social constructivism, to Piagetian constructivism.

Social constructivism, however, was critiqued as downplaying psychological processes in learning. This lead to efforts to create a more robust theory by adding psychological elements from Piagetian constructivism into the sociocultural theories of Vygotsky. As an example, Confrey (1992) relates how Wertsch (1985) "proposes that Piaget should be added into the Vygotskian program" (Confrey 1992: 13) both in the description of "natural" development and in the process of construction of scientific concepts. This, Confrey notes, results "in major changes and contradictions in Vygotsky's program" (1992: 13). This she sees as a general problem with any effort to integrate radical constructivist and sociocultural approaches by simple modifications of each theory.

> *These shifts of attention to include social interaction and cultural influence [in radical constructivism] imply deep reconceptualization of theory and methodology. An integrated theory will need to seek to reshape both theories to allow for both intra-cognitive and inter-cognitive activity. (Confrey 1992: 28)*

Instead, Confrey proposes that an alternative theory is needed, and she lists a number of possible characteristics of such a theory. In the published version of her text, she does not go further than presenting her list. Intriguingly, however, Lerman (1996) in summarising her paper, claims she "argues for a new approach that incorporates Maturana and Varela's (1986) theory of autopoiesis" (Lerman 1996:141). It may be that she argued this when the paper was presented but not in the printed version.

Later, enactivist researchers in mathematics education elaborated the ways in which enactivism in fact offers a new approach that addresses both the individual and the social.

> *In sum, then, cognition does not occur in minds and brains, but in the possibility for (shared) action. Enactivism thus embraces the insights of constructivism, but does not privilege the individual as the truth-determining authority. Similarly, enactivists are able to appreciate the social constructivist's concern for the transcendent (i.e., beyond the individual) nature of knowledge, but do not frame collective knowledge in opposition to subjective knowing. Truth and collective knowledge, for the enactivist, exist and consist in the possibility for joint or shared action—and that, necessarily, is larger than the solitary cognizing agent. Enactivism thus offers a way of bringing these discourses into conversation; for example, constructivism's subject and social constructivism's collective are regarded as self-similar forms. (Davis 1996:192-193)*

While the debates between those taking psychological approaches and those taking sociocultural approaches have largely died down, this is not due to the adoption of enactivism as an alternative theory, as (perhaps) proposed by Confrey. Rather, most mathematics educators have adopted something like the eclectic approach of Cobb and Bauersfeld, using elements of sociocultural and constructivist theories without being too concerned about contradictions and coherence.

It might be asked why enactivism has not been more widely adopted, if it does offer a way to address both the individual and the social in mathematics education. The difficult writing style of Maturana may be a factor, as might be the range of alternative conceptions of enactivism offered by others (e.g., Di Paolo, Rohde and De Jaegher 2010). No one has yet managed to capture the full complexity of enactivism as a biological theory of cognition in language that is more accessible than Maturana's. The confusion around autopoiesis and social systems (explored below) may also be a factor. Many researchers in mathematics education would like to be able to use a common frame to describe cognition in living systems and social systems, and this has resulted in a number of approaches being taken that build on enactivist ideas while not treating enactivism as the full framework desired. Examples include Steinbring's adoption of Luhmann's version of autopoiesis, and Davis' embracing of complexity theory, which is more obviously applicable to complex dynamic systems of any order.

Is enactivism in mathematics education sufficient?

Within mathematics education the critiques of enactivism, and theoretical frameworks in general, have more often been about their sufficiency to address the phenomena of interest to mathematics educators than about their internal consistency. However, the two issues are related. As Confrey's comments above indicate there is a concern that any theory that is sufficient to address both individual learning and social interactions will be incoherent. In this section I will focus on the question of whether enactivism is sufficient within mathematics education, and this discussion will set the stage for a consideration of whether enactivism is coherent.

First it will be necessary to consider the nature of theories in mathematics education from two perspectives. Theories in mathematics education are examples of what Bernstein (2000) calls "horizontal knowledge structures" and they occur both as "grand" theories and as local theories. This establishes the context for the acceptance of enactivism as a theory and its limits when applied to mathematics education. I will then address critiques from within mathematics education concerning the sufficiency of enactivism in describing social systems.

The knowledge structure of mathematics education

As Lerman (2000) points out, education has what Bernstein (2000) calls a "horizontal knowledge structure." This means that new theories in education tend to establish new research domains with their own language. They don't replace other domains, as occurs in vertical knowledge structures like physics, where heliocentrism replaced geocentrism rather than establishing a new research domain alongside it. In mathematics education a new theory like embodied mathematics is not expected to replace other theories, taking over their research domains and transforming them, but rather to establish a new research domain. While proponents of any particular theory would like to think that their theory has the potential to replace others, due to the horizontality of the knowledge structure this does not normally occur.

Furthermore, education in general has a weak grammar; its theories are not able to produce unambiguous descriptions of phenomena. Objects of study in education, such as cognition, learning. knowledge, and emotion, cannot be defined in the way objects of study in physics can be. They are instead described within the frame of reference of a theory. In order to learn what "cognition" means in radical constructivism or embodied cognition "one needs to learn the language of radical constructivism or embodied cognition" (Lerman 2000: 101).

Lerman claims that theories in mathematics education are incommensurable, in principle.

> Where a constructivist might interpret a classroom transcript in terms of the possible knowledge construction of the individual participants, viewing the researcher's account as itself a construction (Steffe and Thompson 2000), someone using socio-cultural theory might draw on notions of a zone of proximal development. Constructivists might find that describing learning as an induction into mathematics, as taking on board concepts that are on the intersubjective plane, incoherent in terms of the theory they are using (and

> a similar description of the reverse can of course be given). In this sense, these parallel discourses are incommensurable. (Lerman 2010: 102)

Grand and local theories

In mathematics education empirical research always takes place with two kinds of theories in play, grand theories and local or intermediate theories, frameworks and models (Ruthven, Laborde, Leach & Tiberghien 2009).

> "Grand theories" [are] theories general in scope and correspondingly abstract in form; notably theories of human development and learning, of the epistemology of the discipline, or of the process of instruction. (Ruthven, Laborde, Leach & Tiberghien 2009: 330)

Grand theories apply not only within mathematics education but to a much wider domain, and within mathematics education a grand theory is expected to address all phenomena of interest. Radical constructivism, sociocultural theory and enactivism are examples of grand theories (Simon 2013). They are expected to be useful in describing any phenomenon of interest in mathematics teaching and learning, either directly, or in some cases by providing a reasoned argument that the phenomenon in question does not exist. *Transmission of knowledge* is an example of a phenomenon that radical constructivism and enactivism would address by questioning and reframing rather than addressing directly.

Local theories, on the other hand, are applicable only within mathematics education, and usually only within a small domain of mathematics education. For example, a theory describing how the principles of arithmetic might be abstracted to become theorems of algebra would be a local theory.

If enactivism is a grand theory then it must address all phenomena of interest to mathematics educators. This includes individual learning, accounting for known phenomena such as the importance of physical materials in learning mathematics, the transition to abstract thinking, and the role of language. It also includes social interactions, including phenomena related to teacher student interactions, student student interactions, and interactions mediated by objects and symbols. Finally it includes the behaviours of social systems.

Apparent insufficiencies of enactivism for mathematics education

Enactivism has been critiqued as not dealing with social interactions, which are undoubtedly important phenomena in mathematics education. For example, Ernest (2010) comments:

> What both enactivism and radical constructivism appear to share is the subordination of the social or the interpersonal dimension, and indeed the existence of other persons to constructions and perceived regularities in the experienced environment. The knowers' own body might be a given, albeit emergent, but other persons' bodies and overall beings are not. Ironically, language, which is the primary seat of metaphor, is the quintessential social construction. But language, like other persons, seems to be removed and exterior to the primary sources of knowledge of the enactive self in these perspectives. (Ernest 2010: 43)

Ernest, however, seems to be referring not to enactivism, but to embodied mathematics. The claim that language and other persons are not central to enactivism is clearly wrong. In fact, recall that the first mathematics educators to use concepts from enactivism, radical constructivists, were interested precisely in ways to refer to social phenomena. This seems to have been forgotten. Perhaps this is a consequence of the way enactivist ideas have been adopted in mathematics education, usually as isolated concepts, grafted onto other theoretical frames. Radical constructivists have adopted the concept of a consensual domain, and the concept of embodied cognition has been employed by researchers interested in gesture, but neither group has actually adopted enactivism as a theoretical frame. This does not mean that enactivism itself is insufficient, however, only that the way it has been employed by radical constructivist and embodied mathematics researchers is insufficient.

Social systems

Ernest's critique above does point out that it is essential to address social aspects of learning in mathematics education. This must be done at two levels, the interpersonal and the social. The interpersonal level includes the social interactions and language use of teacher and students. The social level concerns the behaviour and function of social systems, such as schools, nations and cultures. While enactivism undeniably addresses the social interactions and language use of living systems, it could be critiqued as being unable to address the functioning of non-living social systems. Before responding to this critique it is worth recalling two approaches related to enactivism that have directly addressed social systems. Steinbring's use of Luhmann's sociological theory was described above. Here I will describe Davis's use of complexity theory.

Davis and colleagues (Davis & Sumara 1997, 2006; Davis & Simmt 2003) use complexity theory to extend enactivism to social systems, as well as other complex systems.

> Enactivism understands the individual to be part of—that is, embedded in and a subsystem to—a series of increasingly complex systems with integrities of their own, including classroom groupings, schools, communities, cultures, humanity

> and the biosphere. The notion of “embodied knowledge” extends to bodies much larger than our own. (Davis 1996:193)

This approach extends the types of systems it can address, at the cost of treating them all as complex systems. Complex systems have many important properties, but far fewer than the living systems that enactivism focusses on. This limits the body of concepts that can be brought into the analysis of phenomena in mathematics education. It also risks extending concepts applicable to autopoietic systems, like embodiment, to complex systems in general, as is done in the quote above. There is no reason to believe that this can be done without diluting or destroying the concepts. “Using a concept outside its proper context of application means committing a double fault: the concept will work properly neither in the original nor the new domain” (Maturana & Poerksen 2007: 70)

Enactivism has as its main focus living systems, and so it is much more restricted than complexity theory. However, there is no need to begin with the common properties of all complex systems in an effort to address social phenomena. Maturana himself showed how enactivism ideas can be extended to social systems, and more importantly which ideas can be extended to social systems. Maturana rejected the idea of extending ideas such as autopoiesis, that had been developed in relation to living systems, to social systems without first establishing what the characteristics of social systems are. Hence he begins a paper presented to a symposium on “the theory of autopoietic systems as a new foundation of the social sciences” (Maturana 1980b) by describing what he sees as the organisation of a social system:

> I propose that a collection of interacting living systems that, in the realization of their autopoiesis through the actual operation of their properties as autopoietic unities, constitute a system that as a network of interactions and relations operates with respect to them as a medium in which they realize their autopoiesis while integrating it, is indistinguishable, from a natural social systems and is, in fact, one such system. (Maturana 1980b:12)

This definition of social organisation has not been adopted by mathematics educators and is only referred to briefly by a few radical constructivists. It is also problematic (see below) but it cannot be said that enactivism lacks a way to describe social systems.

Enactivism provides a grand theory that is sufficient to address both the individual and the social in mathematics education. It does so without juxtaposing incompatible frameworks or limiting itself to over general descriptions. But it will not become the dominant grand theory in mathematics education, simply because of the way theories become dominant in domains with horizontal knowledge structures. And there remain aspects of mathematics education that enactivism does not address, most notably the nature and growth of mathematics itself. Other theories must be used to address this aspect.

With the ability to address a wide range of phenomena with a single framework comes the risk of incoherence, and in the next section I will turn to the question of coherence, using my case study of mathematics education research to focus the discussion around the nature of autopoietic systems and social systems.

Is Enactivism Coherent in General?

My case study of enactivism in mathematics education suggests that a possible source of incoherencies is the description of systems using inappropriate terminology. There is a need to be able to describe social systems in particular, and because social systems are like living systems in some ways, there have been efforts to apply Maturana's and Varela's concepts to them. However, unless care is taken to establish the nature of social systems first, there is a danger of misapplying enactivist concepts. Maturana and Varela describe features of living systems, without usually indicating which of the features arise only because the systems are living, and which might apply also to non-living systems. To clarify my discussion of these points I will begin by reviewing the properties of autopoietic systems, which are often given in an abbreviated form that makes it too easy to over generalise the concept. I will then propose a nesting of types of systems and locate living systems, cognitive systems and social systems in it. Through a discussion of the properties of these types of systems, I will locate possible sources of incoherence in enactivism and address them.

Autopoietic systems

It is helpful to be precise about how exactly Maturana and Varela characterise autopoietic systems. Autopoietic systems have a number of properties, all of which must be present for them to be autopoietic. These are the following.

> A. Autopoietic systems are self-producing. They create their own components. "The relations that characterize autopoiesis are relations of *productions* of components" (Varela 1979: 54).
>
> B. Autopoietic systems are embodied. They create a boundary between themselves and everything else. "It is a defining feature of an autopoietic system that it should specify its own boundaries" (Varela 1979: 54).
>
> C. Autopoietic systems are self-organising. The processes or inter-actions between components are organised into a recursive network

that (re)generates itself. “Autopoietic systems … have their own organization as the critical fundamental variable that they actively maintain constant.” (Maturana 1975: 318)

D. Autopoietic systems are composite unities. They are distinguished both as entities and also as networks of interacting components.

E. Autopoietic systems are interactionally open. “Every system will maintain endless interactions with the environment which will impinge and perturb it. If this were not so, we could not even distinguish it.” (Varela & Goguen 1978: 294, original emphasis removed).

F. Autopoietic systems are mechanistic. Their “organization is specifiable only in terms of relations between processes generated by the interactions of components, and not by spatial relations between these components.” (Varela, Maturana & Uribe 1974: 188). Mechanistic systems are structure determined. “A structure determined system is a system such that all that takes place in it, or happens to it at any instant, is determined by its structure at that instant” (Maturana 2002:15).

Properties A and B distinguish autopoietic systems from other autonomous systems, and hence are especially important. However, they are not by themselves sufficient, and focussing on those two properties would mean ignoring important properties that autopoietic systems have in common with other self-organising systems.

Properties C and D together define *organisational closure,* the distinguishing property of autonomous systems.

> An organizationally closed unity is defined as a composite unity by a network of interactions of components that (i) through their interactions recursively regenerate the network of interactions that produced them, and (ii) realize the network as a unity in the space in which the components exist by constituting and specifying the unity’s boundaries as a cleavage from the background (Varela 1981:15)

The distinction between properties A and B, and properties C and D is not always understood, and this results in the misapplication of the label ‘autopoietic’ to autonomous systems that are not autopoietic. “The distinction between autopoiesis as proper to the unitary character of living organisms in the physical space, and autonomy as a general phenomenon applicable in other spaces of interactions, has been consistently confused and left unclarified” (Varela 1981:14). Autonomous systems have properties C and D but not properties A and B.

Property A (self-production) and property C (self-organisation) both concern the generative capacity of the network of interactions in the system. Self-production concerns the production of the system's components. Self-organisation concerns the creation and maintenance of the network of interactions between those components. Neither implies the other. Autocatalytic systems create their own components, and so they are self-producing, but they do not maintain the organisation that allows them to do so, so they are not self-organising (Maturana & Varela 1980: 94). Varela (1979) describes the immune system as an example of an autonomous system, i.e., a system with properties C and D, but its components are produced outside the system so it is not autopoietic.

Property B (embodiment) and property D (unity) both concern the boundary between the system and its medium. However embodiment refers to the production by the system of its boundary which is made of components of the system, while unity refers to the way the system is perceived by an observer, as being both a unity as well as a network of interacting components. The immune system is seen as a system, as a unity, as well as being seen as composed of components that interact. But it does not produce a boundary in the space of its components, which are cells. "The immune system defines a boundary not in a topological sense, but rather in a space of molecular configurations, by specifying what shapes can enter into the ongoing interactions of the system at every point in time" (Varela 1981:18).

The boundary between the system and its medium is also related to property E (interactional openness). A boundary both marks the extent of a system, as well as providing the means by which it interacts with its medium. The importance of interactions between the system and its medium through its boundary is indicated by this recent definition of autopoiesis: "A system is autopoietic if: (a) it has a semi-permeable boundary, (b) the boundary is produced from within the system, and (c) it encompasses reactions that regenerate the components of the system." (Varela 2000, in Bourgine & Stewart 2004: 329). A system that did not interact with its medium would be unobservable, and in addition, would not last long as it would have no way to import energy to offset entropy.

Property F (mechanistic) places autopoietic systems among dynamic systems, which are defined by recursive properties rather than by geometric characteristics, which can also give rise to emergent properties (as is the case in crystals). This property is also a reminder that autopoiesis is a non-vitalist description of life. Being alive comes from recursive properties of the system rather than the presence of a vital spark or substance.

Nesting of types of systems

Of the properties of autopoietic systems listed above, property E must be the most general. If a unity is not open to interactions with an observer, it cannot be observed. All the other properties can only apply to observable unities. This is because properties are not properties of the unity but properties of the observer's interactions with it.

> The basic cognitive operation that we perform as observers is the operation of distinction. By means of this operation we specify a unity as an entity distinct from a background, characterize both unity and background with the properties with which this operation endows them, and specify their separability. (Maturana & Varela 1980: xix)

Property D distinguishes composite unities from simple unities.

> A unity thus specified [by an operation of distinction] is a simple unity that defines through its properties the space in which it exists and the phenomenal domain which it may generate in its interactions with other unities. If we recursively apply the operation of distinction to a unity, so that we distinguish components in it; we respecify it as a composite unity that exists in the space that its components define because it is through the specified properties of its components that we observers distinguish it. Yet we can always treat a composite unity as a simple unity that does not exist in the space of its components, but which exists in a space that it defines through the properties that characterize it as a simple unity. (Maturana & Varela 1980: xix)

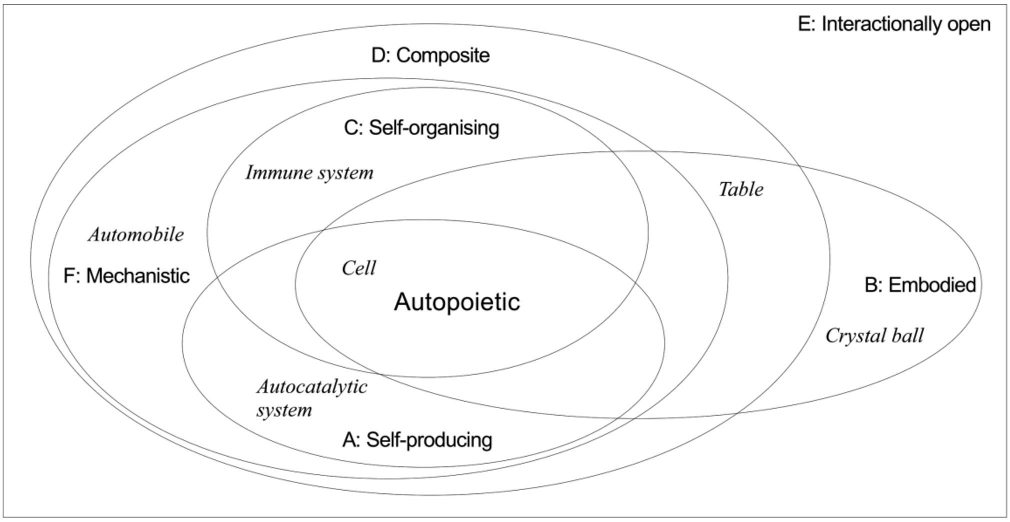
Some unities can only be observed as simple unities. Which unities are simple depends on the observer of course. In my case, given a sphere of clear crystal I observe it as a simple unity. Other observers say that the crystal ball has components, atoms in a particular configuration, but I do not observe those components, so it is a simple unity for me. Composite unities can be observed in two ways, either as a simple unity or as a set of components. Varela (1979) calls these two ways of observing the 'behavioural view' and the 'recursive view' respectively. Maturana and Varela use the word 'system' to refer only to composite unities observed recursively, which is a narrower usage than is common is systems theory generally, and which can lead to seeming incoherencies.

Property B can apply to both simple unities and composite unities. I observe the crystal ball as having a topological boundary between it and not-it. I can also observe through a microscope the topological boundary between a cell and its medium. Because the cell is a component entity, I can observe its boundary in two ways, either as an edge between it and not-it, or as a component of the cell itself.

Within composite unities we can distinguish between static unities and dynamic or mechanistic unities. A static unity is distinguished by the spatial relations between its components. As an observer I distinguish a table as a simple unity by the interactions I can have with it, and as a composite unity by the way its parts are put together. A dynamic or mechanistic unity is distinguished by the "relations between processes generated by the interactions of components" (Varela, Maturana & Uribe 1974: 188). The spacial relations between the components of a mechanistic unity are not fixed, but its components interact in ways that define its organisation.

Finally, I discussed above autonomous systems and self-producing systems. These are overlapping subtypes of mechanistic systems. Autopoietic systems are located in the intersection of autonomous systems, self-producing systems and embodied unities (See Figure 1).

Figure 1: Nested types of systems



Having identified the key properties of autopoietic systems and having shown how they are nested, I will now turn to some specific systems of interest: living systems, cognitive systems and social systems, and discuss where they fit into the nesting of properties.

Living systems

A question that has been discussed a great length is whether or not the categories 'living system' and 'autopoietic system' are identical. Maturana and Varela initially coined the term 'autopoiesis' in order to characterise living systems, and claimed "autopoiesis is necessary and sufficient to characterize the organization of living systems" (1980:82). This suggests that all autopoietic systems are living systems and all living systems are autopoietic. However, as soon as the word was defined, it began to be applied to systems that are not usually thought of as living, including social systems. Maturana and Varela may have contributed to this confusion themselves by describing a computer model as autopoietic (Varela, Maturana & Uribe 1974, Varela 1979).

If autopoiesis is not sufficient to characterise living systems, what else is required? Maturana clarifies "An autopoietic system that exists in physical space is a living system (or, more correctly, the physical space is the space that the components of living systems specify and in which they exist)" (1978a: 36). Computer models do not exist in physical space, and so are not alive. What exactly is physical space? "The physical space is defined by components that can be determined by operations that characterize them in terms of properties such as masses, forces, accelerations, distances, fields, etc." (Maturana & Varela 1980: 112).

Another characterisation is that living systems are autopoietic systems whose components are molecules. In fact, this requirement of having molecular components came before the word 'autopoiesis' was coined.

> At the beginning of the year 1964 I began to say that living systems were constituted as unities or discrete entities as circular closed dynamics of molecular productions open to the flow of molecules through them in which everything could change except their closed circular dynamics of molecular productions. (Maturana 2002: 8)

Maturana goes on to say "I also claim that autopoiesis occurs only in the molecular domain" (Maturana 2002: 8). At one point he did think it "possible that autopoietic systems could exist in domains different from the molecular one" (Maturana 2002: 14) however, he later came to see the molecular domain as having unique properties necessary for autopoiesis.

> Molecules through their interactions give rise to molecules and dynamic systems of molecular productions, in diffuse and localized processes that constitute discrete entities. I think that due to this peculiarity of the molecular domain this is the only domain in which autopoietic systems can take place as discrete singular systems that operate through thermal agitation and dynamic architecture. (Maturana 2002: 8)

This seems to be a claim that could be empirically tested. In the cybernetic tradition, whether or not a system is autopoietic ought to be a matter of the system's organisation, independent of the nature of its components. "The actual nature of the components, and the particular properties that these may possess other than those participating in the interactions and transformations which constitute the system, are irrelevant and can be any" (Maturana & Varela 1980: 77). This leaves open the possibility that there might be non-living, non-molecular autopoietic systems. Of course, Maturana has every right to narrow the meaning of 'autopoiesis' to apply only to living, molecular systems, making autopoiesis a matter not only of organisation but also of the type of components involved. This would only mean that another word would be required to describe systems that have the same organisation as living, molecular systems, but are made of non-molecular components. While I sympathise with Maturana's fears that the concept of autopoiesis could become useless if it becomes ill defined or over-applied, there seems to me to be some value in describing living systems as *molecular* autopoietic systems, leaving open the possibility that autopoietic systems could be created from other components. I would agree with Varela:

> The relations that characterize autopoiesis are relations of productions of components. … Given this notion of production of components, it follows that the cases of autopoiesis we can actually exhibit, such as living systems or the example described in Varela et al. (1974), have as a criteria of distinction a topological boundary, and the processes that define them occur in a physical-like space, actual or simulated in a computer. (Varela 1981:15)

Computer models can be autopoietic in a "physical-like space." While Varela's original model has been critiqued (on the basis that it included a component, a catalyst, that it cannot produce, Bourgine & Stewart 2004) efforts continue to produce computer models that have all the properties of an autopoietic system, within the space they define. In other words, I would argue that living systems are located entirely in the intersection of autonomous systems, self-producing systems and embodied unities (see Figure 2) but that they do not fill it. There could be autopoietic systems that are non-living.

Cognitive systems

Enactivism is a theory of cognition, in which cognition is seen as a property of all living systems, which are defined as autopoietic systems.

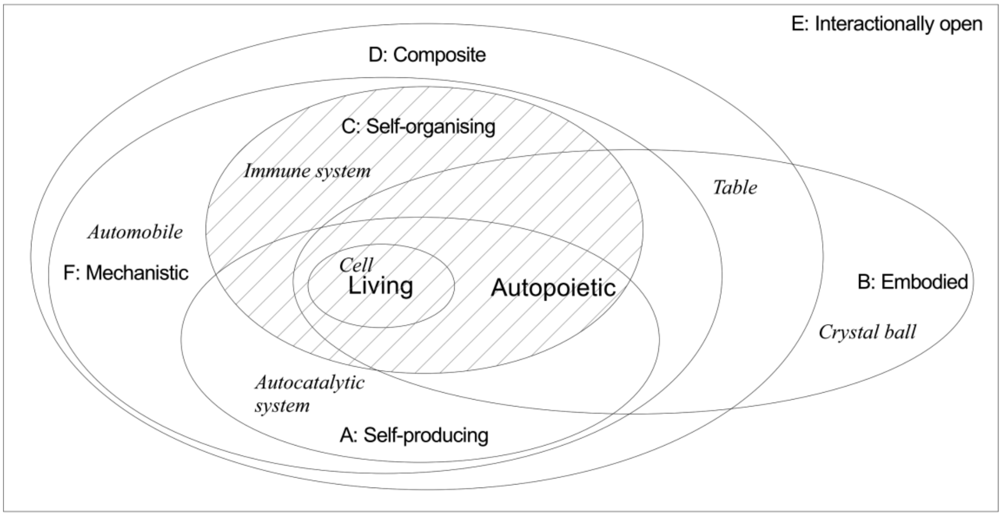
> A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself, and the process of cognition is the actual (inductive) acting or behaving in the domain. Living systems are cognitive systems, and living as a process is a process of cognition. This statement is valid for all organisms, with and without a nervous system. (Maturana 1980a: 13).

Succinctly, cognition is "effective behavior in a medium" (Maturana 1978b: 37). What types of systems can behave effectively in a medium? Living systems certainly can, and it is the embedding of learning and cognition in a general study of life that makes enactivism distinctive. But living systems are not the only systems that define a domain of interactions in which they can act with relevance to the maintenance of themselves. All that is required is that the system engage in maintaining itself, and that true of autonomous systems in general. "The mechanisms of identity of an autonomous system correlate with the establishment of cognitive interactions with its environment" (Varela 1979: 211). Varela gives as two examples the nervous system and the immune system, but he also discusses cognitive social systems, including transient interactions like conversations.

> Every autonomous structure will exhibit a cognitive domain and behave as a separate, distinct aggregate. Such autonomous units can be constituted by any processes capable of engaging in organizational closure. whether molecular interactions, managerial manipulations, or conversational participation.... I am saying, then, that whenever we engage in social interactions that we label as dialogue or conversation, these constitute autonomous aggregates, which exhibit all the properties of other autonomous units. (Varela 1979: 269)

This broadening of the meaning of cognition is useful in educational research, where a focus on organisms with nervous systems, especially people, can obscure fundamental issues such as the role of structural coupling in learning. Research on learning and cognition is not restricted to human learning and cognition, or even the cognition of living systems. Cognitive systems are located in and coincide with autonomous systems (see Figure 2).

Figure 2: Cognitive and living systems



Social systems

If there is an area where enactivism is incoherent, it may well be in the treatment of social systems. This is perhaps not surprising, as Maturana and Varela specialised in the study of living systems, and their comments on social systems did not rest on the same level of expertise. In addition, if one assumes that the components of social systems are living systems, as they did, and therefore that the components of human social systems are human beings, then one is limited in the type of observations that are possible.

> Nothing prevents the observer himself from being part of the process of specifying the system, not only by describing it, but by being one link in the network of processes that defines the system. This situation is peculiar in that the describer cannot step outside of the unity to consider its boundaries and environment simultaneously, but it is associated with the unity's functioning always as a determining component. Such situations, to which most of the autonomous social systems belong, are characterized by a dynamics in which the very description of the system makes the system different. At each stage, the observer relates to the system through an understanding which modifies his relationship to the system. This is, properly speaking, the hermeneutic circle of interpretation-action, on which all human activity is based. (Varela 1981: 16)

Human beings observing human social systems are limited in two ways. First, they cannot take a behavioural view on the system, seeing it a simple unity, stepping "outside of the unity to consider its boundaries and environment". Human observers of human social systems are always components of the system or its environment and can only observe it with a recursive view, focussed on its components and the interactions between them. Second, describing a social system is a way of interacting with its components, in a way that describing a cell is not: "The very description of the system makes the system different." Given these challenges it is not surprising that Maturana and Varela had some difficulties describing social systems.

However, there are some points about which they are clear and consistent. First, that integrating social systems into enactivism must begin by understanding social phenomena independently of enactivist concepts and terminology. Second, that the components of social systems are living systems. And third, that social systems are not autopoietic.

Maturana begins his 1980 essay 'Man and society' by asking "What is a social system?" (1980b: 11) and giving as the criterion for judging an answer to this question comparison to "the same phenomena that a natural social system appears to generate in its operation" (1980b: 11). In other words, before proposing an answer to the question "What is a social system?" it is necessary to observe social phenomena. Maturana saw this as a fundamental problem in Luhmann's use of the concept of autopoiesis.

> I suggest that we start with the question of the characteristics of social phenomena. The concept of society historically precedes the idea of the autopoiesis of living systems. Society was the primary subject of debate; autopoiesis and social systems came much later. It follows, therefore, that we should first deal with all the relevant phenomena appearing in the analyses of society and only afterwards ask ourselves whether they may be elucidated more precisely in terms of the concept of autopoiesis. (Maturana & Poerksen 2007: 70)

Another disagreement Maturana had with Luhmann concerned the components of social systems. Luhmann proposed that the components of a social system are communications and that human beings form a part of the medium in which the social system exists. Maturana rejected this position: “When we speak about social systems in our everyday life, however, we naturally have in mind all the individuals with their peculiar properties” (Maturana & Poerksen 2007: 71). “A social system can only be integrated by living systems” (Maturana 1980: 13). Maturana’s clearest definition of a social system starts with living systems:

> [A social system is] a collection of interacting living systems that, in the realization of their autopoiesis through the actual operation of their properties as autopoietic unities, constitute a system that as a network of interactions and relations operates with respect to them as a medium in which they realize their autopoiesis while integrating it. (Maturana 1980b: 11-12)

Although Maturana and Varela could not agree completely on how to treat social systems (Maturana & Varela 1980: xxiv), Varela also assumes that the components of human social systems are human beings, as indicated in his remark about autonomous social systems quoted above (Varela 1981: 16).

Both Maturana and Varela are clear that they do not see social systems as autopoietic. Social systems can be autonomous (i.e., they have properties C-F) but social systems do not have boundaries and do not produce their components.

> There have been proposals suggesting that certain human systems, such as an institution, should be understood as autopoietic (Beer 1975; Zeleny and Pierre 1976). From what I have said I believe that these proposals are category mistakes: they confuse autopoiesis with autonomy. (Varela 1981: 15)

However, Maturana occasionally makes comments that makes this point less clear.

> Just imagine for a moment a social system that is, in actual fact, functioning autopoietically. It would be an autopoietic system of the third order, itself composed of autopoietic systems of the second order. This would entail that every single process taking place within this system would necessarily be subservient to the maintenance of the autopoiesis of the whole. Consequently, the individuals with all their peculiarities and diverse forms of self-presentation would vanish. They would have to subordinate themselves to the maintenance of autopoiesis. Their fate is of no further relevance. They must conform in order to preserve the identity of the system. This kind of negation of the individual is among the characteristics of totalitarian systems. Stalin, therefore, forced party members who did not share his outlook to give up their positions so as not to endanger the cohesion and the unity of the party. In a democratic form of communal life, however, individuals are of central relevance and, in fact, indispensable. Their properties create the unique character of a social system. (Maturana & Poerksen 2007: 72)

For Maturana, a social system is “a medium in which [living systems] realize their autopoiesis” (Maturana 1980b: 12), which means he must reject any system that interferes with the autopoiesis of the living systems in it. But as he notes, totalitarian systems do precisely this. By Maturana's definition, totalitarian systems are not social systems. "A social system that forbids and even principally excludes complaint and protest is not a social system. It is a system of tyranny." (Maturana & Poerksen 2007: 72). Given that Maturana developed his definition of a social system while in exile from Pinochet's Chile, he may have had reasons to characterise social systems in the way he did. However, it is not only totalitarian systems that are excluded by Maturana's definition.

> A person who works for a given society and who cannot stop working for it without risking the loss of his autopoiesis because he has no other means of survival outside this work-relation, is under social abuse. Example: In a capitalistic economic system a worker is not a member of the productive society through which he earns his living and, therefore, only works for it. If, under these circumstances, there is no employment with respect to his abilities, and if he has no other independent means of survival, he is under social abuse. Such a person cannot enter into a work-agreement on terms generated by the fundamental equality that permits cooperation, and must surrender his autonomy as a human being in order to survive. (Maturana 1980b: 18)

Armies, police departments, fire departments and other organisations that subsume their members' autopoiesis to the goals of the organisation are also excluded by Maturana's definition of a social system. However, Maturana states that "In general any organism, and in particular any human being, can be simultaneously a member of many social systems, such as a family, a club, an army, a political party, a religion or a nation, and can operate in one or other without necessarily being in internal contradiction" (Maturana & Varela 1980: xxviii). It is not at all clear that armies, religions and nations are systems that support the autopoiesis of their component human beings. And armies, religions and nations routinely restrict complaint and protest. Perhaps Maturana did not mean to include all religions and nations as social systems, but it is hard to imagine any army that would fit his definition of a social system.

A further aspect of Maturana's definition of a social system is the role of love in constituting a human social system.

> What determines the constitution of a social system are the recurrent interactions of the same autopoietic systems. In other words, any biological stabilization of the structures of the interacting organisms that results in the recurrence of their interactions, may generate a social system. Among human beings the basic stabilizing factor in the constitution of a social system is the phenomenon of love, the seeing of the other as a partner in some or all the dimensions of living. (Maturana & Varela 1980: xxvi)

It may be that Maturana has in mind an ideal human society, rather than any actual human society.

> A human society in which to see all human beings as equivalent to oneself, and to love them, is operationally legitimate without demanding from them a larger surrender of individuality and autonomy than the measure that one is willing to accept for oneself while integrating it as an observer, is a product of human art, that is, an artificial society that admits change and accepts every human being as not dispensable. (Maturana & Varela 1980: xxix)

If we accept Maturana's definition of social system, then we require another word for those composite unities that most people consider to be social systems: families, clubs, armies, political parties, religions, nations, etc. Either that or we need another definition of 'social system' in enactivist terms. Because Maturana's definition of a social system is problematic, and was never accepted by Varela, enactivism lacks a coherent definition of 'social system' derived from primary sources. This has left the field open to many proposals of alternative definitions, from Beer (1980) to Zeleny (Zeleny & Hufford 1991). This host of alternatives makes it impossible to place social systems definitively in the nesting of types of systems depicted in Figure 1. Social systems may or may not be autopoietic, but there seems to be a general agreement that they are at least autonomous, and so have many interesting properties.

Conclusion

To conclude I will reiterate a few of the main points I have made above. First, in mathematics education research enactivism offers a 'grand theory' that can be brought to bear on most of the phenomena of interest to mathematics educators. It has particular strength in describing interactions between cognitive systems, including human beings, human conversations and larger human social systems. Much remains to be done in exploring the potential of enactivism for social cognition. Second, some apparent incoherencies of enactivism come from the adoption, in mathematics education but also in other fields, of parts of enactivism which are then grafted onto incompatible theories. This opens up enactivism to critiques from both within mathematics education and outside the field. Most strongly, in my opinion, theories of cognition that claims to be enactivist, but rely only on the philosophical arguments introduced in *The Embodied Mind* rather than the biological arguments presented in *Autopoiesis and Cognition,* leave enactivism open to philosophical critiques. This is one reason for my insistence on referring primarily to Maturana's work in defining enactivist concepts. Third, and finally, a source of incoherence is the lack of a generally agreed upon definition of a social system. There is no reason why a suitable definition cannot be found, and I suspect replacing Maturana's 'love' with something like a shared emotional orientation (Maturana 1988) would be sufficient, but this remains to be done.

References

Barnes, J. M. 1994. *Human-computer enaction,* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Bateson, G. 2000. *Steps to an ecology of mind.* Chicago: University of Chicago Press.

Bateson, G. 1979. *Mind and nature: A necessary unity*. New York: Dutton.

Bateson, G. 1972. *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. Lanham, MD: Jason Aronson.

Bauersfeld, H. 1980. Hidden dimensions in the so-called reality of a mathematics classroom. *Educational studies in mathematics*, 11(1): 23–41.

Beer, S. 1980. Preface to Autopoiesis. H. Maturana & F. Varela, eds. *Autopoiesis and cognition: The realization of the living*: 63–72. Dordecht: D. Reidel Publishing Co.

Begg, A. 2013. Interpreting enactivism for learning and teaching. *Education sciences & society,* 4(1): 81–96.

Begg, A. 2002. Enactivism and some implications for education: A personal perspective. *Vinculum*, 39(2): 4–12.

Begg, A. 1999. Enactivism and mathematics education. In J. Truran & K. Truran, eds. *Making the difference: Proceedings of the twenty-second annual conference of the Mathematics Education Research Group of Australasia*: 68–75. Adelaide, South Australia:MERGA.

Bernstein, B. 2000. *Pedagogy, symbolic control and identity: Theory, research, critique.* Revised edition. Lanham, MD: Rowman & Littlefield.

Blackmore, D., Cluett, G. & Reid, D. 1996. Three perspectives on problem-based mathematics learning. E. Jakubowski, D. Watkins, & H. Biske, eds. *Proceedings of the Eighteenth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education,* Volume 1: 258–262. Columbus OH: ERIC Clearinghouse for Science Mathematics, and Environmental Education.

Bourgine, P. & Stewart, J. 2004. Autopoiesis and cognition. *Artificial life*, 10(3): 327–345.

Brousseau, G. 1997. *Theory of didactical situations in mathematics: Didactique des mathématiques, 1970–1990*. Eds & Trans. N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield. Dordrecht: Kluwer.

Brown, L. & Coles, A. 2010. Mathematics teacher and mathematics teacher educator change: Insight through theoretical perspectives. *Journal of mathematics teacher education,* 13(5): 375–382.

Brown, L. & Coles, A. 2008. *Hearing silence: Steps to teaching mathematics.* Cambridge UK: Black Apollo Press.

Brown, L. & Coles, A. 2000. Complex decision making in the classroom: the teacher as an intuitive practitioner. T. Atkinson & G. Claxton, eds. *The intuitive practitioner:*

On the value of not always knowing what one is doing: 165–181. Buckingham, UK: Open University Press.

Brown, L. & Coles, A. 1997. The story of Sarah: seeing the general in the particular? E. Pehkonen, ed. *Proceedings of the twenty-first annual conference of the International Group for the Psychology of Mathematics Education,* Volume 2: 113–120. Lahti, Finland: PME 21.

Brown, L. & Reid, D. 2006. Embodied cognition: Somatic markers, purposes and emotional orientations. *Educational studies in mathematics,* 63(2): 179–192.

Brown, L. & Reid, D. 2004. Emotional orientations and somatic markers: Implications for mathematics education. M.J. Høines & A.B. Fuglestad, eds. *Proceedings of the Twenty-eighth conference of the International Group for the Psychology of Mathematics Education,* Volume 1: 123–126. Bergen, Norway: PME

Brown, L. & Reid, D. 2003. Observing the unseeable: Somatic markers related to students' and teachers' decision-making. Williams, J. ed. *Proceedings of the British Society for Research into Learning Mathematics,* 23(2): 19–24.

Brown, L. & Reid, D. 2002. Somatic markers: Teachers' decision-making and students' emotioning in mathematics classrooms. *Proceedings of the British Society for Research in the Learning of Mathematics,* 22(3): 7–12.

Brown, L., Reid, D. A. & Coles, A. 2003. Seeing patterns: Somatic markers in teachers' decision-making and students' reasoning in mathematics classrooms. *Research in mathematics education*, 5(1): 17–30. DOI:10.1080/14794800008520112

Campbell, S. & Dawson, A. J. 1995. Learning as embodied action. R. Sutherland, & J. Mason, eds. *Exploiting mental imagery with computers in mathematics education*: 233–249. Heidelberg: Springer.

Campbell, S. 2010. Embodied minds and dancing brains: New opportunities for research in mathematics education. B. Sriraman & L. English (Eds), *Theories of mathematics education. Seeking new frontiers*: 309–331. Heidelberg: Springer.

Chevallard, Y. 1999. L'analyse de pratiques professorales dans la théorie anthropologique du didactique. *Recherches en didactique des mathématiques,* 19(2): 221–266.

Cobb, P. 1986. Contexts, goals, beliefs, and learning mathematics. *For the learning of mathematics*, 6(2): 2–9.

Cobb, P. 1985. Two children's anticipations, beliefs, and motivations. *Educational studies in mathematics,* 16(2): 111–126.

Coles, A. 2011. *Metacommunication and Listening: An enactivist study of patterns of communication in classrooms and teacher meetings in one secondary mathematics department in the UK*. Doctoral dissertation. Bristol UK: University of Bristol.

Confrey, J. 1992, February. How compatible are radical constructivism, social-cultural approaches and social constructivism? Paper presented at "Alternative Epistemologies in Education," Athens, Georgia. Source: https://gismodb.-fi.ncsu.edu/files/resources/How_Compatible.pdf, 27.11.2013.

Damasio, A. 1994. *Descartes' error: Emotion, reason, and the human brain.* New York: Avon Books/ G. P. Putnam's Sons.

Davis, A. B., Sumara, D. & Kieren, T. 1996. Cognition, co-emergence, curriculum. *Journal of curriculum studies*, 28(2): 151–169.

Davis, B. & Simmt, E. 2003. Understanding learning systems: Mathematics education and complexity science. *Journal for research in mathematics education*, 34(2): 137–167.

Davis, B. & Sumara, D. 2006. *Complexity and education: Inquiries into learning, teaching, and research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Davis, B. & Sumara, D. 1997. Cognition, complexity, and teacher education. *Harvard educational review,* 67(1): 105–126.

Davis, B. 2004. *Inventions of teaching: A genealogy.* Mahwah, NJ: Lawrence Erlbaum.

Davis, B. 1997. Listening for differences: An evolving conception of mathematics teaching. *Journal for research in mathematics education*, 28(3): 355–376.

Davis, B. 1996. *Teaching mathematics: Toward a sound alternative.* New York: Garland Publishing.

Davis, B. 1995. Why teach mathematics? Mathematics education and enactivist theory. *For the learning of mathematics*, 15(2): 2–9.

Davis, B. 1994. *Listening to reason: An inquiry into mathematics teaching.* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Edwards, L. & Núñez, R. (1995). Cognitive science and mathematics education: A non-objectivist view. L. Meira & D. Carraher, eds. *Proceedings of the nineteenth annual conference of the International Group for the Psychology of Mathematics Education*, Volume 2: 240–247. Recife: PME.

Ernest, P. 2010. Reflections on theories of learning. B. Sriraman & L. English, eds. *Theories of mathematics education. Seeking new frontiers*: 39–47. Heidelberg: Springer.

Ferrara, F. 2003. Metaphors as vehicles of knowledge: An exploratory analysis. N. Pateman, B. Dougherty & J. Zilliox, eds. *Proceedings of the twenty-seventh annual conference of the International Group for the Psychology of Mathematics Education,* Volume 2: 373–380. Honolulu, HI: PME.

Gordon Calvert, L. 1999. *Mathematical conversations within the practice of mathematics.* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Gordon Calvert, L., Kieren, T., Reid, D. & Simmt, E. 1995. Thinking otherwise: Enactivism, mathematics and education. Paper presented at the meeting of the Canadian Society for the Study of Education, Montreal.

Hadamard, J. 1945. *The mathematician's mind: The psychology of invention in the mathematical field*. Princeton, NJ: Princeton University Press.

Johnson, M. 1987. *The body in the mind: The bodily basis of reason and imagination.* Chicago: University of Chicago Press.

Kenny, V. 2007. Distinguishing Ernst von Glasersfeld's radical constructivism from Humberto Maturana's 'Radical Realism'. *Constructivist foundations*, 2(2–3): 58–64.

Kieran, C., Forman, E. & Sfard, A. 2001. Guest Editorial: Learning discourse: Sociocultural approaches to research in mathematics education. *Educational studies in mathematics*, 46(1): 1–12.

Kieren, T. 1994. Orthogonal reflections on computer microworlds, constructivism, play and mathematical understanding. *Journal of research in childhood education*, 8(2): 132–141.

Kieren, T. & Simmt, E. 2009. Response to Fazio and Gallagher: Brought forth in bringing forth. The inter-actions and products of a collective learning system. *Complicity: An international journal of complexity and education*, 6(2): 20-28.

Kieren, T., Gordon Calvert, L., Reid, D. & Simmt, E. 1996. Occasioning: Learning in the interaction. Presented at the Annual Meeting of the American Educational Research Association, New York.

Kieren, T., Gordon Calvert, L., Reid, D. & Simmt, E. 1995. An enactivist research approach to mathematical activity: Understanding, reasoning, and beliefs. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Kilpatrick, J. 1992. A history of research in mathematics education. D. Grouws, ed. *Handbook of research on mathematics teaching and learning*: 3–38. New York: Simon & Schuster Macmillan.

Konold, C. & Johnson, D. 1991. Philosophical and psychological aspects of constructivism. In L. Steffe, ed. *Epistemological foundations of mathematical experience*: 1–13. New York: Springer.

Krainer, K. 1993. Powerful tasks: A contribution to a high level of acting and reflecting in mathematics instruction. *Educational studies in mathematics*, 24(1): 65–93.

Lakatos, I. 1976. *Proofs and Refutations*. Princeton: Princeton University Press.

Lakoff, G. & Núñez, R. 2000. *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York: Basic Books,

Lakoff, G. 1987. *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lerman, S. 2010. Theories of mathematics education: Is plurality a problem? B. Sriraman & L. English, eds. *Theories of mathematics education. Seeking new frontiers*: 99–109. Heidelberg: Springer.

Lerman, S. 1996. Intersubjectivity in mathematics learning: A challenge to the radical constructivist paradigm? *Journal for research in mathematics education*, 27(2): 133–150.

Lozano, M-D. 2004. *Characterising algebraic learning: An enactivist longitudinal study.* Unpublished doctoral dissertation, University of Bristol.

Luhmann, N. 1997. *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.

Luhmann, N. 1995. *Social systems*. Stanford: Stanford University Press.

Luhmann, N. 1986. The Autopoiesis of Social Systems. F. Geyer and J. Van d. Zeuwen, eds. *Sociocybernetic paradoxes: Observation, control and evolution of self-steering systems*: 172–192. London: Sage.

Mason, J. & Johnston-Wilder, S., eds. 2004. *Fundamental constructs in mathematics education*. New York: RoutledgeFalmer.

Mason, R. 1997. *Learning algebra personally.* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Maturana, H. & Poerksen, B. 2007. Autopoiesis and social theory: A conversation. *Journal of sociocybernetics*, 5(1-2): 68–73.

Maturana, H. & Varela, F. 1987. *The tree of knowledge: The biological roots of human understanding*. Boston: New Science Library/Shambhala Publications.

Maturana, H. & Varela, F. 1980. *Autopoiesis and cognition: The realization of the living.* Dordecht: D. Reidel Publishing Co.

Maturana, H. 2002. Autopoiesis, structural coupling and cognition: A history of these and other notions in the biology of cognition. *Cybernetics & human knowing*, 9(3–4): 5–34.

Maturana, H. 1988. Reality: The search for objectivity or the quest for a compelling argument. *The Irish journal of psychology*, 9: 25–82.

Maturana, H. 1987. Everything said is said by an observer. W. Thompson, ed. *Gaia: A way of knowing*: 65–82. Hudson, NY: Lindisfarne Press.

Maturana, H. 1980b. Man and society. F. Benseler, P. Hejl & W. Köck, eds. *Autopoiesis, communication, and society: The theory of autopoietic systems in the social sciences*: 11–32. Frankfurt: Campus Verlag.

Maturana, H. 1980a. Biology of cognition. H. Maturana & F. Varela, eds. *Autopoiesis and cognition: The realization of the living*: 5–58. Dordecht: D. Reidel Publishing Co.

Maturana, H. 1978b. Cognition. P. Hejl, W. Köck & G. Roth, eds. *Wahrnehmung und Kommunikation*: 29-49. Frankfurt: Peter Lang.

Maturana, H. 1978a. Biology of language: The epistemology of reality. G. Miller & E. Lenneberg eds. *Psychology and biology of language and thought: Essays in honor of Eric Lenneberg*: 27–63. New York: Academic Press.

Maturana, H. 1975. The organization of the living: A theory of the living organization, *International Journal of Man-Machine Studies,* 7: 313-332.

Nicol, C. & Lerman, S. with assistance from Joop van Dormolen, Carolyn Kieran, Gerard Vergnaud, Kath Hart and Heinrich Bauersfeld. 2008. *A brief history of the International Group for the Psychology of Mathematics Education (PME) Source: http://www.icmihistory.unito.it/pme.php*, 29.11.2013.

Núñez, R. & Edwards, L. 1995. Embodied cognition and the psychology of mathematics education: A discussion gorup. L. Meira & D. Carraher, eds. *Proceedings of the nineteenth annual conference of the International Group for the Psychology of Mathematics Education*, Volume 1: 184. Recife: PME.

Núñez, R., Edwards, L. & Matos, J. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational studies in mathematics*, 39(1–3): 45–65.

Oehrtman, M. C. 2003. Strong and weak metaphors for limits.B. Dougherty & J. Zilliox, eds. *Proceedings of the twenty-seventh annual conference of the International Group for the Psychology of Mathematics Education,* Volume 3: 397–404. Honolulu, HI: PME.

Pirie, S. & Kieren, T. 1989. A recursive theory of mathematical understanding. *For the learning of mathematics*, 9(3): 7–11.

Poincaré, H. 1914. *Science and method*. DoverPublications.com.

Radford, L., Edwards, L. & Arzarello, F. 2009. Introduction: beyond words. *Educational studies in mathematics,* 70(2): 91–95.

Reid, D. & Brown, L. 1999. Fork in the road. *For the learning of mathematics,* 19(3): 15–22.

Reid, D. & Drodge, E. 2000. Embodied cognition and the mathematical emotional orientation. *Mathematical thinking and learning,* 2(4): 249–267.

Reid, D. 2002b. Elements in accepting an explanation. *Journal of mathematical behavior,* 20(4): 527–547.

Reid, D. 2002a. Conjectures and refutations in grade 5 mathematics. *Journal for research in mathematics education,* 33(1): 5–29

Reid, D. 1999. Needing to explain: The mathematical emotional orientation. Orit Zaslavsky. ed. *Proceedings of the twentieth-third annual conference of the International Group for the Psychology of Mathematics Education*, Volume 4: 105–112. Haifa, Israel: PME.

Reid, D. 1996. Enactivism as a methodology. L. Puig & A Gutiérrez, eds. *Proceedings of the twentieth annual conference of the International Group for the Psychology of Mathematics Education*, Volume 4: 203–210. Valencia, Spain: PME.

Reid, D. 1995. *The need to prove.* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Reid, D., Brown, L. & Coles, A. 2001. *Observing systems.* Working session at the twentieth-fifth annual conference of the International Group for the Psychology of Mathematics Education, Utrecht, July 2001.

Ruthven, K., Laborde, C., Leach, J. & Tiberghien, A. 2009. Design tools in didactical research: Instrumenting the epistemological and cognitive aspects of the design of teaching sequences. *Educational researcher*, 38(5): 329–342.

Samson, D. & Schäfer, M. 2010. Beyond constructivism: Enactivism as a theoretical lens in the context of figural pattern generalisation. V. Mudaly, ed. *Proceedings of the eighteenth annual meeting of the Southern African Association for Research in Mathematics, Science and Technology Education (SAARMSTE)*, Book of Abstracts: 21. Durban, SA: University of Kwazulu-Natal.

Samson, D. & Schäfer, M. 2011. Figural apprehension, enactivism and knowledge objectification: An exploration of figural pattern generalisation. *For the learning of mathematics*, 31(1): 37–43.

Samson, D. & Schäfer, M. 2012. The case for manipulatives: Enabling "ways of seeing". In D. Nampota & N. Kazima, eds. *Proceedings of the twentieth annual meeting of the Southern African Association for Research in Mathematics, Science and Technology Education (SAARMSTE),* Volume 1: Long Papers: 220–232. Lilongwe, Malawi: SAARMSTE.

Simmt, E. 2000. *Mathematics knowing in action: A fully embodied interpretation.* Unpublished doctoral dissertation. Edmonton, Alberta: University of Alberta.

Simon, M. A. 2013. The need for theories of conceptual learning and teaching of mathematics. K. Leatham, ed. *Vital directions for mathematics education research*: 95–118. New York: Springer.

Skemp, R. 1971. *The psychology of learning mathematics*. Baltimore, MD: Penguin Books.

Sriraman, B. & English, L. (Eds). 2010. *Theories of mathematics education. Seeking new frontiers*. Heidelberg: Springer.

Steinbring, H. 2005. Analyzing mathematical teaching-learning situations: The interplay of communicational and epistemological constraints. *Educational studies in mathematics*, 59(1–3): 313–324.

Steinbring, H. 1999. Mathematical interaction as an autopoietic system: Social and epistemological interrelations. I. Schwank, ed. *European Research in Mathematics Education, Proceedings of the First Conference of the European Society for Research in Mathematics Education* Volume 1-2: 387– 400. Osnabrück: Forschungsinstitut für Mathematikdidaktik.

Thompson, P. 2008. Conceptual analysis of mathematical ideas: Some spadework at the foundation of mathematics education. O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sépulveda, eds. *Proceedings of the Thirty-second annual meeting of the International Group for the Psychology of Mathematics Education*, Volume 1: 45–64. Morélia, Mexico: PME.

Varela, F. (1981). Autonomy and autopoiesis. In G. Roth, & H. Schwegler, eds. *Self-organizing systems: an interdisciplinary approach*:14-24. New York: Campus Verlag.

Varela, F., Thompson, E. & Rosch, E. 1991. *The embodied mind: Cognitive science and human experience.* Cambridge MA: MIT Press.

von Glasersfeld, E. 1991. Distinguishing the observer: An attempt at interpreting Maturana. *Methodologia*, V (8): 57–68 von Glasersfeld, E. 1989. Cognition, construction of knowledge, and teaching. *Synthese*, 80(1): 121–140.

Wertsch, J. 1985. *Vygotsky and the social formation of mind.* Cambridge MA: Harvard University Press.

Zeleny, M. & Hufford, K. 1991. All autopoietic systems must be social systems (living implies autopoietic. But, autopoietic does not imply living): An application of autopoietic criteria in systems analysis. *Journal of social and biological structures*, 14(3): 311–332.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0008

Enactivism and the New Teleology: Reconciling the Warring Camps

Ralph D. Ellis
Clark Atlanta University
rellis[]cau.edu
Received December 2013; accepted September 2014; published Autumn 2014.

Abstract

Enactivism has the potential to provide a sense of teleology in purpose-directed action, but without violating the principles of efficient causation. Action can be distinguished from mere reaction by virtue of the fact that some systems are self-organizing. Self-organization in the brain is reflected in neural plasticity, and also in the primacy of motivational processes that initiate the release of neurotransmitters necessary for mental and conscious functions, and which guide selective attention processes. But in order to flesh out the enactivist approach in a way that is plausible and not merely an epiphenomenon, it is necessary to confront the problem of causal closure in a serious way. Atoms and molecules in the brain do not violate the normal causal principles that govern them in other contexts. The theory of self-organizing dynamical systems must be developed in a way that is compatible with causal closure rather than contradicting it.

Keywords: Enactive; self-organization; dynamical systems; neural plasticity; causal closure.

Introduction

Clinical psychologists tend to view human behavior in teleological terms, while experimentalists (including many of those same clinicians when adopting a more theoretical stance) view all causation as efficient, with human actions appearing not so much as self-initiated actions, but as reactions to stimuli. The enactive approach to cognition can go a long way toward reconciling these different perspectives. When enactivism is grounded in a sober analysis of causal relations and neural mechanisms (as in Kauffman 2003; Monod 1971), it can make room for organismic purposes that drive both action and the nature of re-actions. Multiple realizability and neural plasticity are major components of this analysis. In such processes, the same organizational activity can make use of different sets of micro-level components, which the organism actively appropriates and replaces as needed, insofar as possible given environmental conditions and the organism's particular needs.

However, the structure of causal analyses within enactivist and self-organizational theory must be taken more seriously than often is done. Vague talk of "higher level" processes that "constrain" causal relations at a lower level can be misleading or even meaningless. The problem of causal closure (Kim 1992, 1998) must be confronted rather than merely talked around. When the causal analyses are done in a careful way, and integrated with the motivational systems of the brain that orchestrate self-initiated and self-energized actions, explanations can remain consistent with normal efficient causal processes that are by no means violated by the atoms and molecules that make up the nervous system. Yet at the same time, enactivist cognitive theory combined with the neurological phenomena of neural plasticity and the serious working out of self-organizational causal theory can make room for a meaningful distinction between action and a merely complicated sequence of re-actions.

This paper will attempt to provide an overview of the path that can lead from self-organizational causal theory, through mechanisms of neural plasticity in the nervous system that serve organismic purposes, to the primacy of motivated selective anticipation in perception, to the enactivist manifestations of these mechanisms in cognitive processes. I will refer here to examples of the timing and interactions of brain mechanisms in perception, imagination, and motivated attention direction (discussed more fully in Ellis 1995, 2005; Ellis and Newton 2010) that illustrate the primacy of purposeful processes of the anticipatory (Freeman 2001) and searching (Panksepp 1998, 2012) functions of the motivational brain. These self-energizing and anticipatory systems lead to views of cognitive functions that fit nicely with enactivist approaches such as those of Noë (2006) and Thompson (2007).

But equally important, I will also discuss the philosophical analyses that are needed to reconcile self-organizational and enactive processes (the "autopoiesis" emphasized by Thompson) with the problem of causal closure of physical systems. More is needed for this reconciliation than merely multiple realization; we need a concept of a *particular kind* of multiply realizable system in which the overall pattern of organization is of such a nature that it shows a tendency to appropriate, replace, and reproduce the components of its own definitive patterns of activity by virtue of the organized way it exchanges energy and materials with the environment (Newton 2000). Such a theory has been extensively developed by Stuart Kauffman (1993) and Scott Kelso (1995).

Science and subjectivity

"Psychology as a science" is a strange expression. It means eliminating from the study of the psyche—the subject of experience—everything that does not present itself in the form of an object of experience. Above all, this means eliminating teleology, because modern science explains things from the standpoint of mechanistic causal relations. Rolls (1999), for example, explains emotion in these terms, and the account of perceptual processing is standardly framed in terms of sequences of efficient causes in the brain (for example, see any standard textbook such as H.R. Schiffman 2001). John Bickle (2008) expresses the philosophical rationale for rejecting teleology in favor of reductionist causal accounts. When I decide to raise my hand, and then do so, we must explain the causes of motion for a measurable and observable physical mass. Modern science therefore must explain the raising of the hand in terms of objectively observable physical causes. And this means in terms of the electro-chemical reactions in the nervous system, which can be reduced to a series of micro-level causal sequences. These causal sequences occur at a level at which things obey the laws of classical Newtonian mechanics, and thus are causally closed. If physical antecedents are necessary and sufficient as an explanation of them, then the conscious correlates of those physical antecedents can play no causal role, unless we completely reduce consciousness to its physical components. This equation of the conscious with its physical components is now a standard solution to this problem, endorsed for example by Bickle as the only reasonable way to make physical and mental causation compatible with each other. If consciousness is reducible to a physical sequence of micro-level events, then it can have the same causal powers as that sequence of events.

And even if we do opt for a strict identity between consciousness and its physical correlates, the past century of work on psychophysical monism still seems to leave us with only mechanistic causes—but with some of them correlating with mental events—and thus without any real teleological causal relations, but only a subjective impression of one. Instead of raising my hand *for the purpose* of reaching the coffee cup, strictly speaking, I am raising it because of neurophysiological events that correlate with the *image* of the coffee cup, which then physically triggers a hand movement, all of which in principle is explainable by physical mechanisms at the level of molecules and chemical reactions.

If the physical components correlating with the thought "I'd like to raise my hand" completely explains the physical raising of the hand, then the idea that the conscious thought caused the hand to raise has only the *appearance* of a teleological process. The relation between the mere appearance of a teleological process and the underlying reality of a mechanistic process is the same as when birds *appear* to migrate "for the purpose" of finding a better climate to spend the winter. In reality (in the standard mechanistic scientific account) there ultimately must be micro-level causal explanations for what causes them to follow the patterns of flight that they follow. The real explanation for why they migrate in those patterns has no more to do with the birds' achieving a "purpose" than the causal explanation of a giraffe's long neck is that it stretches "in order to" reach the leaves in high branches. To be sure, there are natural selection processes at work, but these in turn are ultimately reducible to a series of mechanistic cause-and-effect relations. Natural selection, far from embracing teleology, actually explains it away.

In no sense is the giraffe's long neck *mechanically caused* by the giraffe's "having a purpose," nor is the birds' pattern of migration mechanically caused by the purpose of the migration. So if all causation is mechanistic, the "purpose" achieved by the birds and the giraffe play no causal role, because this having-a-purpose does not play the causal role that the underlying mechanistic antecedents play.

In theory, it has seemed since the inception of "psychology as a science" that teleological explanations of human behavior and consciousness must work the same as with the giraffe and the birds. It subjectively *appears* that our having a purpose is what causes us to raise our hand, but the real causal work is done by physical antecedents that preceded any subjective impression of "having a purpose," and which are entirely sufficient to explain the resulting behavior without reference to any feeling of teleology. In reality, in the dominant streams of purely empirical psychology, our human behavior is as mechanistic as any other causal relation. While these mechanisms seem purposeful from the subjective perspective, objectively they result from push-pull dynamics. The past is always what causes the future. The same problem presents itself for other apparently purposeful processes in nature, such as the remarkable shunt mechanisms in the metabolism of nutrients in animal systems and many other non-conscious contexts. Our focus here, however, is the need for something resembling a kind of teleology in conscious processes. We will see later that the self-organizational theories that reconcile mental with physical causation are also the same ones that reconcile non-conscious purposeful processes with their micro-level efficient causal mechanisms.

Although many would like to avoid the problems of classical Newtonian mechanics by explaining consciousness at the quantum level (as in Jibu and Yasue 1994; Globus 1992, 2003), in my view this move still cannot avoid the problem of mental causation and other related contradictions between the subjective experience of teleology and the mechanistic explanations of its neurophysiological correlates. Even if quantum behavior does not fit the causally closed picture of Newtonian physics, the fact remains that conscious processes have been correlated thousands and thousands of times with causal mechanisms that *do* occur at a level where causal closure still obtains. We cannot avoid these correlations by descending to the quantum level of explanation. That move does not erase the correlations.

Classical Newtonian causal relations are causally closed (as emphasized, for example, by Kim 1992, 1993, 1998). If some combination of conditions are sufficient to produce an outcome, then no other condition can be necessary to produce that same outcome. Conversely, if some combination of conditions are necessary to produce an outcome, then no other condition can be sufficient to produce that same outcome. The result is that teleological explanations of why and how I raise my hand can make no reference to my subjective decision to raise it, because the raising of the hand has been entirely accounted for by micro-level physical conditions—unless, of course, I completely reduce the conscious event to a physical one, in which case we still are left with no teleological causal relations, only mechanistic ones.

Adding further to the reductionism of this schema is the passive receiving model of perception, which leads to stimulus-response explanations of all psychological phenomena, including affective and conscious ones. If I am quietly sitting, without any light impinging on my retina, I do not see any object. When the light hits the retina, I see an object. Many psychologists are content to infer that the difference between seeing and not seeing must be caused by the introduction of the light hitting my retina—i.e., the stimulus. My consciousness of the stimulus is then the subsequent product of this causal dynamic. Consciousness in effect is *caused by* physical actions done *to* the body by external objects—not by actions initiated by the organism which would then use or adapt to the external objects. So consciousness must be an epiphenomenon, a mere side effect of the activity of stimuli as they hit my nervous system and in turn lead to cause and effect mechanisms in it. From this standpoint, consciousness is always construed as the last event in a series of causes and effects.

There is of course one easy way out of the standard mental causation problem. If consciousness is identical to a physical process, then whatever is caused by that physical process is also caused by the corresponding conscious act (as in Smart 1970). In that case, even a conscious choice could still become an intermediate cause in a chain, since it would be identical with some of the intermediate physical events in the chain. To be sure, consciousness would only be an intermediate step in a chain of mechanistic causes in that case, but at least it would play a causal role.

Up to this point, the dominant schools of empirical psychology have taken this intermediate causal role of consciousness to mean that we have a subjective reaction to a stimulus input, and a conditioned (mechanistic) response, of which we can form a subsequent subjective impression—as in behaviorism. Or this view can be qualified by allowing innate configurations of brain matter, hormones, and neurotransmitters to partially determine *how* we respond to a given stimulus—as in the traditional drive theories such as those originally developed by Hull (1952) and Spence (1956; see also Spence and Frith 1999), and currently handed down to us by Rolls (1999), LeDoux (1996) and others.

But of course such a literal psycho-physical identity theory presents a plethora of other problems. For example, there is the "hard problem" pointed out by Chalmers (1995), which is essentially a special case of Levin's (1983) "explanatory gap" problem. In Chalmers' formulation, the problem is that, if we can explain everything about a physical event, and if the physical event is identical with a conscious event, then this would mean that we have completely explained the conscious event. But the problem is that something would still be left unexplained: Namely, after we have explained what caused physical event XYZ to occur, we still would not have explained why an event like XYZ should have the property of consciousness, whereas other physical events subject to the same kinds of explanations do not have the property of consciousness. To explain what caused XYZ to occur does not explain why XYZ has the property of consciousness. Those are two different questions. The second one is what Chalmers means by the "hard problem" of consciousness. He explicitly distinguishes this from the "easy problems" of consciousness. Explaining that consciousness always correlates with XYZ is easy. And explaining what caused XYZ to occur is also easy, compared with the "hard problem." Yet one often hears presentations of solutions to these easy problems accompanied by the claim that they are solving, or will lead to the solution of the "hard problem." This way of speaking betrays a misunderstanding of what Chalmers means by the "hard problem."

I use Chalmers' "hard problem" as only one of many problems that are created by a literal psycho-physical identity theory. As far as teleology is concerned, psycho-physical identity remains at odds with it even if we completely accept psycho-physical identity. Consciousness is still caused by an interaction of mechanistic physical events. And here again, the standard application of empirical psychology to this model has been in effect to use stimulus response explanations, combined with some assumptions about innate physiological predispositions. The result is that "action" is explained as a series of micro-level *re*-actions. We are left with no distinction between action and reaction.

Clinical psychologists have long wished that there *could* be such a thing as teleology in human processes, because our subjective reporting of what we feel to be the purposes of our actions and motivations is largely experienced as teleological. I do not experience myself as deciding to listen to Tchaikovsky's Sixth Symphony because some classical or operant conditioning history caused something to trigger me to respond to some stimulus in my environment in some way. My experience is that I first *felt* a certain way, and because I felt that way, I wanted to listen to some music appropriate to that feeling. Rather than being caused by the music, the feeling pre-exists and then uses the music. I engage in voluntary actions needed for access to the right kind of music—driving to a concert hall or playing a disk on a stereo.

The enactive approaches to psychology offer the kind of teleology that is needed here. As the name implies, enactivism begins with the idea that there is such a thing as action—not merely a sequence of *re*-actions. As is now well known, new scientific work on the theory of self-organization in certain kinds of physical systems promises to allow the whole organism to act from its own self-initiated motivations, and not merely to react or to display a complicated system of reactions (see Kauffman 1993; Monod 1971; Newton 1996, 2000; Ellis and Newton 2010; Ellis 2005).

The distinction between action and reaction, construed scientifically, can be grounded in the idea of complex dynamical systems, which I take as synonymous with self-organizing systems. As Newton (2000) defines it, a dynamical system is a thermodynamic system that maintains the continuity of its functional patterns while exchanging energy and materials with the environment. Such a system acts as a whole to seek out, select, appropriate, rearrange, and reproduce the micro-constituents needed to keep its ongoing pattern going with continuity. As Thompson (2007) suggests, the pattern uses its parts to maintain the continuation of the pattern and, within the limits of its situation, replaces the parts as needed for this purpose.

We must be very careful to avoid facile explanations of the relationship between the mechanistic level and the self-organizational level of causal analysis. Locutions like "the higher level of organization *constrains* what can happen at the lower level" (as in Thelen and Smith 1994; Globus 2003) may be satisfying for some purposes, but they do not answer the crucial questions about causation. In classical mechanical causation, which the neurons and neurotransmitters of the brain are observed to obey quite consistently, the cause (or combination of causes) is necessary and sufficient for its effects under the given background conditions. This is as true for turning on an electric light as for shunt mechanisms in biological organisms. So if saying that the higher level "constrains" what the lower level can do simply means that lower level causal relations are sufficient only after we have assumed that certain background conditions are in place (for example, a good bulb in the light socket and a complete electrical circuit), then we still have not distinguished between linear and complex dynamical causal systems. We need to be careful to give a coherent explanation as to how a higher-level of organization "constraining" the lower level is any different from any other causal situation. How are self-organizing systems different from non-self-organizing ones?

In my view, the difference hinges on the fact that, in self-organizing systems, the system as whole displays a tendency for its various components to be rearranged when needed to provide the background conditions that are assumed by a needed micro-level causal relation (Kauffman 1993; Monod 1971).

We do not need for this overall plasticity to be devoid of previous mechanistic explanations for how it initially achieved the plastic condition—for example, the micro-level reactions that occurred during embryonic development (as both Kauffman and Monod point out). The point is that very early in the etiology of a self-organizing system, it acquires a tendency to rearrange its own components in such a way that the structural integrity of the whole is maintained by constantly rearranging the background conditions needed for the given micro-level linear reaction to occur (see Ellis and Newton 2010). This is the sense in which the overall causal situation is different from simply a combination of linear reactions. The whole system seeks out, appropriates, rearranges and replaces its own parts as needed to maintain the pattern of the whole or to change the pattern in ways called for by the needs of the whole system. We see this in many examples of neural plasticity, such as in stroke recovery (Restak 1984: 256; Springer and Deutch 1981: 173-212), in recovery from memory loss from head injuries (Restak 1984:. 360ff; Wrightson 1989), and in drug addiction (in which neurotransmitter receptors change their size and number to avoid imbalances resulting from the substances that are oversupplied by the drug user—for example, see Valenstein 2002, Chapter 5).

How can such a dynamical systems approach avoid the problems of previous mechanistic explanations? Is the dynamical system merely an epiphenomenon of the entire series of mechanistic interactions of its components, and thus in reality only a series of reactions after all? Does the system as a whole actually violate causal closure in ways that would contradict a huge mass of empirical evidence suggesting that there are systematic correlations between physical cause-effect sequences and conscious events? Or is the action of the dynamical system somehow "compatible with" the mechanistic explanation of the behavior of its micro-level constituents? Those are pressing problems that enactivism must face if it is to ground itself in action in a meaningful sense, rather than only an epiphenomenal appearance of action.

The new approach has arisen partly from a rejection of some of the assumptions of the stimulus-response and passive-reaction models of the past century These assumptions, as we have seen, led to the relegating of consciousness to an epiphenomenal role in philosophy and science. They ultimately stem from the assumption that the reality which ultimately must explain all causation is thoroughly "bottom-up," with the real causal work done at an atomistic-reductionist level; that representational conscious activities (thoughts and perceptions) are clearly distinguishable from non-representational ones (feelings and emotions); and, perhaps most important, that all reality is fundamentally reactive and passive rather than active—i.e., that nothing does anything unless caused to do it by some external force acting on it, that there is no such thing as a pattern of activity which organizes its own substrata rather than the other way around. In short, for "modernist" metaphysics (as postmodernists like Globus, 1992, 2003 call it), there was no important or non-arbitrary distinction between non-living things and living ones (i.e., those which appropriate, rearrange, and reproduce the needed substrata in order to maintain a higher-order pattern of activity); yet the difference between conscious beings and non-conscious ones (e.g., computers) would need to hinge crucially on this distinction.

What is needed for a teleological yet scientific analysis of intentional actions

In the modernist framework as applied to psychology and cognitive neuroscience, consciousness is a causal result of a stimulus input which then gets processed in the brain. Perceptual consciousness is thus assumed to result from stimulation of the occipital lobe and V4 visual areas, as a result in turn of a perceptual object's stimulating the nervous system. Thomas Natsoulas (1994) calls this approach an "appendage" theory of consciousness—the idea that consciousness is a byproduct of a cause and effect mechanism but does not have the power to act as a cause itself. But there are problems with trying to interpret perceptual consciousness in this way.

For example, it has often been observed that occipital activation in response to incoming visual data often results in no perceptual consciousness. Additionally, it is necessary that the parietal and frontal lobes must also play an active role (Farah 1989; Luria 1980; Posner 1980, 1990; Posner and Petersen 1990); and it is well established that this activation of the parietal and frontal lobes is not simply caused by the earlier activation of the occipital cortex (Aurell 1989). As I discussed more extensively elsewhere (Ellis 1995, 2005), the parietal lobe is almost immediately adjacent to the occipital lobe, yet in the process of generation of a conscious perceptual experience, the parietal lobe is not active until about a third of a second after the occipital lobe is activated when a novel stimulus is presented (Runeson 1974; Srebro 1985; McHugh and Bahill 1985). Given that nerve impulses travel at about 100 miles per hour, why should it take a third of a second to travel only a few millimeters?

If the imaginative activity of the parietal lobe were really caused by the nerve impulse which travels to it from the occipital lobe, the impulse should be delivered virtually instantaneously. Whatever is happening during this third of a second that is also needed in order for consciousness of the object to occur cannot be caused by the passive receiving of the nerve impulse to the parietal lobe from the occipital lobe, which in turn receives it from the incoming stimulus. If so, the spreading of activation would be much quicker than a third of a second; it would be virtually instantaneous. Thus it appears that the response is not caused by the stimulus.

This paradox arises only if we assume that the parietal lobe (which is active when we are conscious of visual images) can only be activated as a result of prior occipital activity, which in turn results from prior optic stimulation originating from the environment. But recent research shows that this is not the case. Instead, what happens is that, prior to occipital processing of the visual stimulus, the parietal lobe has already been activated by the frontal lobe (as shown by Aurell 1989, Damasio et al. 2000, and many others), which in turn is activated by the midbrain, which is the focus of emotional-motivational activity as triggered by thalamic arousal by the stimulus only if the stimulus is generally felt as possibly emotionally important for the organism's purposes (Luria 1980; Posner 1990; Damasio 1994). The needs of the organism as a whole must first motivate the asking of questions about what kinds of environmental stimuli might be important for the organism's purposes; at this point, the frontal lobe becomes active. As these questions are formulated with the help of the frontal lobe, the parietal lobe then begins to entertain vague images and/or concepts of the kinds of emotionally important objects that might be present in the environment. If and when this frontal-limbic-parietal activity, once having been developed, finds itself resonating with patterns of activity in the occipital lobe (which reflects sensory stimulation)—only then does perceptual consciousness occur.

The one-third-second delay does not result from any slowing of the incoming nerve impulse as it “travels” from the occipital lobe to the parietal lobe. The parietal lobe (which is active when we are conscious of visual images) is not activated in response to the occipital lobe's activity at all. Instead, the organism must purposely activate the frontal and parietal lobes to “look for” emotionally important categories of objects which the thalamus has already alerted the organism might be relevant, and this “looking for” activity has already begun the forming of visual or conceptual imagery (including proprioceptive and sensorimotor imagery) prior to any occipital activity's having any effect on our perceptual consciousness (since at this point the impulse has not yet “traveled” from the occipital to the parietal lobe). As Thompson (2007), Noë (2006), and Ellis and Newton (2010) formulate this process, the organism is always already engaged in a series of both conscious and non-conscious actions, and then makes an assessment of the environment based on how it seems to facilitate or thwart the process. Neurophysiologically, the prefrontal cortex’s searching function and the hippocampal and amygdala emotional processes are already active long before the occipital perceptual processing occurs (Damasio et al. 1990; Haines et al. 1997).

On the enactivist account of intentionality, the organism must act on its environment in order to be conscious of it; consciousness cannot result from a mere passive reaction to incoming input. On this view, the model of the mind as a passive receiver of causal work done by stimulus inputs and other mechanical computations is backwards in some important respects. The organism must first purposely act, and only then can consciousness of the environment result. It is this fundamental shift in the direction of causation which is now sometimes referred to as the "enactive" view of the mind—a term coined by Varela et. al. (1991/1993). Rather than a stimulus causing a response, it is the response which must occur first, and then act on the incoming afferent signals to produce a stimulus. We might call this enactive approach the current "Copernican revolution" in cognitive theory and neuroscience.

The sensorimotor account of intentionality

The philosopher Natika Newton has elaborated an entire theory of consciousness around the notion that, in effect, we must subliminally use "sensorimotor action images" to imagine the action affordances of objects in order to be consciously aware of them (this careful analysis is developed through a sequence of detailed studies, including Newton 1982, 1989, 1991, 1992, 1993, 1996, 2000). According to Newton, perceptual consciousness is always preceded by an act of imagination. We first imagine ways in which we could possibly act in relation to the object at hand. The action motivated by the action planning process creates an "anticipation" as to possible environmental feedback, and these anticipations constitute mental images. If the anticipations are fulfilled by correlative perceptual input, the result is a percept; if the anticipations are not fulfilled, then the result often is still a state of consciousness, but the state consists only of a mental image of a non-present object or state of affairs. The anticipation, however, must precede the effect of the incoming data from the senses on our perceptual consciousness. Subjunctive ideas about imagined possibilities for action are thus prior to perceptual input, and action planning guides the process of "looking for" instantiations of the subjunctive category (for example, the image) as actually instantiated in the environment. For example, perceptual studies by Runeson (1974), Srebro (1985) and others show that an object previously anticipated is much more readily perceived.

Newton's theory is a clearly articulated example of an enactivist account in which consciousness must play an active role in the context of the experience of humans and other higher animals. The role it plays is similar to the "pragmatic" role that Marcelo Dascal (1987) assigns to it in bringing about the kind of information processing that is accompanied by conscious experience. But if the organism must act in this sense, it is not merely reacting to stimuli, and therefore is not an easy fit to the notion of an epiphenomenon or "appendage" to a basically non-conscious computational process. This aspect of the enactivist account builds from the foundation established by J.J. Gibson (1986) in his notion of "affordances." We understand and identify an object by imagining how it would be possible for our bodies to act in relation to the object. Newton and other recent enactivists use neurological work (for example, Damasio

1994; Posner1990; Farah 1989; Luria 1980; Cytowic 1993), and research on mental imagery (for example, the “mental models” developed by Johnson-Laird and Byrne 1991) to show how action planning grounds our understanding of objects, and ultimately of language, concepts, and logical relations. An infant identifies objects in terms of whether they “afford sucking,” “afford throwing,” etc.

This imagistic account can ground our understanding of logical relations because subjunctive action affordances allow us to imagine what would happen “if” such and such circumstances were to obtain. For example, to anticipate that “If I throw a ball at something it will knock it over,” is very similar to believing that “If I *were* to throw a ball, it *would* knock something over.” Thus anticipations of the future ground our understanding of subjunctives and therefore allow a grasping of abstract logical concepts. In Newton's approach, the key to this “foundation of understanding” is the process of action planning. This theory is consistent with a host of neurophysiological findings—for example, that the brain mechanisms underlying abstract thought are very similar to those underlying action-planning in the context of body movement (Ito 1993; Damasio 1994).

The crucial point for our purposes here is that, in the enactivist approach, the modernist biases mentioned above become questionable and a new paradigm suggests itself. Because the organism must anticipate actions toward its environment in order for consciousness to occur, consciousness is not merely passively caused by incoming stimuli or unconscious computations performed on incoming stimuli. The body's organization of stimuli occurs prior to the reception of the stimuli, and if the body does not actively seek to appropriate and rearrange the physiological substrata for its own desired patterns of conscious activity, this consciousness can never occur.

In a sense, we could think of the relationship between neurons and the pattern of activity that constitutes consciousness as similar to the relationship between the molecules of wood in a door and the sound wave that passes through the door. The wave is composed of a pattern of the movements of the wood particles, but we would not want to “explain” the wave by reference to the door. The door does not cause the wave, but on the contrary, the wave causes the particles in the wooden door to conform to its pattern as it passes through the door.

Correlatively, an enactive approach such as Newton's or Varela's requires a rejection of the epiphenomenalist account of consciousness as merely the tip of an iceberg which consists of unconscious computational brain processes. Instead, consciousness directs much of this activity, and much of it would never take place without the direction of consciousness; yet it is important that consciousness itself is embodied—not in computational cerebral processes, but rather in emotional and motivational activities of the whole organism.

It is the emotionally motivated process of action planning that directs the focus of conscious attention, not a computer-like computational process.

For the same reasons, enactivism must reject the modernist assumption that representational states (thoughts and perceptions) are caused by perceptual inputs. In fact, in the enactivist account, representational processes are not even clearly distinguishable from non-representational ones (feelings and emotions). Emotion and motivation direct the action imagery that grounds the experience of the action affordances of objects. The purposes of the whole embodied organism come into play in directing conscious attention. Much of our rational processing, for the reasons that I sketched earlier, results from what Newton calls "proprioceptive imagery." In proprioceptive imagery, we imagine what it would be like to perform a bodily action, and this in turn grounds the understanding of a subjunctive concept of such a movement and of the environmental objects that afford it. The understanding of "what would happen if…" grounds our learning of which kinds of logical inferences can be relied upon as we go through life. And much of this action imagery can occur on a subliminal or pre-conscious basis. Only in the presence of a strong enough motivating selective attention process and a suitable environmental context (or a sufficiently rich imagining of one), does consciousness accompany the action imagery.

In essence, twentieth century empirical psychology was built on an ontology of the physical world in which everything is fundamentally reactive rather than consisting of patterns of activity which appropriate their atomistic components. When scientists tried to apply this passive ontology to the explanation of consciousness, the result was that no explanation was possible. The easy problems of consciousness could be addressed, but no attempt could even be made at the "hard problem." The next section will consider the advantages of enactivism in making more sense out of the relationship between consciousness and the natural world than was possible in the passive-receiving model of the mind and the epiphenomenalist conceptualization of the relevant causal relations.

How enactivism can confront the "hard problem" head-on

The approach that I have been outlining is equipped to offer a new perspective on Chalmers' (1995) "hard problem" of consciousness. Computationalists had maintained throughout most of the twentieth century that consciousness can be explained either as an epiphenomenon of, or as identical with, a digital computer-like process which uses the hardware of the brain to process its software. The point of Chalmers' "hard problem" is that we can easily imagine such a computational process as occurring in the absence of consciousness. Therefore, some further explanation is required in order to understand why consciousness does in fact accompany such computational processes in certain cases (for example, in human organisms).

The enactive approach we have been discussing avoids this problem because it focuses on elements of consciousness as actually phenomenally experienced that were left out of the traditional computationalist model—elements that lend themselves to being correlated with action-oriented as opposed to passive-receiving physiological substrata. It is not so easy to imagine that this teleological or action-initiating combination of physiological substrata, including emotionally motivated selective attention, imagery, and resulting action, could have occurred unaccompanied by its conscious correlates. For example, consider the enactive model of consciousness that Newton and I described in a co-authored article and subsequent further elaboration in book form (Ellis and Newton 1998, 2010). On the formulation developed there, consciousness requires

> an interested anticipation of possible sensory and proprioceptive input such that the pattern of the subject's interest determines the modality, patterns, and emotional significance of the anticipated input. Specifically, the anticipation takes the form of a sensorimotor, proprioceptive and affective "image" of a state of affairs "looked for" by the subject.... The content of consciousness is vivid to the extent that the activity constitutive of the interest in the future resonates (in terms of holistic patterns of activity) with the activity of incoming (afferent) imagistic data and with activation of memories of past imagistic and conceptual data. (Ellis and Newton 1998: 432)

Subjects can actually experience this effect in perceptual attention experiments. When the subjects are instructed in such a way that they are anticipating an object before it is presented, they perceive it more readily (Corbetta 1990; Pardo et al 1990; Logan 1980; Hanze and Hesse 1993; Legrenzi et al 1993; Rhodes and Tremewan 1993; Lavy and van den Hout 1994). Posner and Rothbart (1992) report that "During vigilant states the posterior attention network can interact more efficiently with the object recognition system of the ventral occipital lobe (96)." This attentional process "increases the probability of being able to detect the relevant signal (97)." Imagining a given object involves a process that is very similar, both phenomenologically and neurophysiologically, to "looking for" an object of that type, even if we are aware that we will not find it in the actual current environment. When I form a mental image of the pink wall as blue, I am executing many of the same brain processes as when I "look for" or "anticipate" blue in the wall, even though I do not find the blue I am "looking for" in this particular wall. As Merleau-Ponty says, "I give ear, or look, in the expectation of a sensation, and suddenly the sensible takes possession of my ear or my gaze, and I surrender a part of my body, even my whole body, to this particular manner of vibrating and filling space known as blue or red" (1962: 212). Later he sums up in this way: "It is necessary to 'look' in order to see" (1962: 232). And "The warmth which I feel when

I read the word 'warm' is not an actual warmth. It is simply my body which prepares itself for heat and which, so to speak, roughs out its outline" (1962: 236). Helmholtz (1962) makes a similar point which is now widely accepted among neurologists: "We let our eyes traverse all the noteworthy points of the object one after another." I.e., the organism must actively search for information in the environment before that information is consciously seen. Vision is active, not passive.

One of the main differences between conscious and non-conscious information processing (as in nuts and bolts computers) is that consciousness involves an emotionally *interested* anticipation of a possible or actual stimulus input (Cytowic 1993). In order to be conscious of a specific input, we must actively and motivatedly "pay attention," as inattentional blindness experiments clearly show (Mack and Rock 1998; Ellis 2005). The action of directing attention, like any other action, must be motivated by the needs of the organism. To be sure, afferent processing involving the occipital lobe is *part* of what produces conscious awareness of an object, but this awareness also requires corticothalamic loops instigated by subcortical motivational activity and the cooperation of frontal-limbic loops with the anterior cingulate (see Bachmann 2000; Lethin 2002, 2005; Posner and Rothbart 1992; Damasio 1994; Farah 1989; Aurell 1989; Luria 1980.) In the enactive approach, the primary organismic need that motivates consciousness of objects is the need to anticipate future data which are considered important for the organism's purposes (Dennett 1996).

The emotionally motivated anticipation of input leads to "imagery" in an enactivist rather than modernist sense. In enactivist terms, an "image" is not a physical replica or "encoding" of the object in the brain, but rather an enacting of the process that one would undergo if perceiving the object—in other words, on my formulation, a sense that we are looking for (or listening for, tasting for, proprioceptively feeling for, etc.) an object or state of affairs in a widely distributed neurophysiological pattern similar to what *would* be executed if we were to actually see, hear, or taste it. This idea is highly consistent with Walter Freeman's (1991) findings, for example that cats activate their olfactory patterns merely in anticipation of the presentation of the smell of fish, before the smell is actually presented. It is also consistent with the now familiar finding that imagining a given action activates most of the same brain processes as actually executing the action. Jeannerod (1994, 1997), for example, explains that when we imagine a movement, we execute all the same brain processes as in actually doing the movement, but at the same time we inhibit the action command at the point when it would feed forward to the body's extremities in order for the action to be carried out.

When we say that consciousness is distinguished from non-conscious processing by being emotionally motivated, this should not be taken as implying that emotions and motivations are necessarily conscious. There are non-conscious yet purposive phenomena in nature, especially in biological organisms. For example, the human organism purposefully does what is necessary to regulate its heartbeat and blood pressure, yet normally is not conscious of doing so. Merleau-Ponty (1941) defines a "purposeful organism" as one which changes, replaces, or readjusts the functioning of its own parts according to what is needed to maintain or enhance the existence and functioning of the whole organism. The purpose-directing role of emotion is not a sufficient condition for consciousness, but in the enactivist account as I have formulated it, emotion and motivation are necessary conditions.

It is true that we can have interested anticipations of the future in this sense without any experienced consciousness. This can occur because, throughout nature, there is purposeful activity without conscious awareness, because self-organization in complex systems can occur in various contexts throughout nature. It is also true that we can process afferent data without consciousness, as in "blindsight" phenomena. And of course there is also holistic processing without consciousness, as in typical holograms. There are also sometimes non-conscious interests alongside of non-conscious data processing, but without consciousness. We can even have non-conscious anticipations of the future, as in operant conditioning, juxtaposed with non-conscious activations of stored information or of present afferent activity, with no consciousness of the process.

So enactivists should be careful not to try to explain consciousness in terms of any *one* of these elements without the needed contextualization in relation to the others—as for example people sometimes speak as if complex dynamical systems could automatically explain what makes something conscious.

Consciousness requires that the interest in the future is neurophysiologically instantiated so as to give rise to an image of a possible future. In the case of perceptual consciousness, this same activity resonates with afferent activity stimulated by input. The interest in the future, the forming of the image, and the processing of the sensory or sensorimotor data must all resonate with each other as the motivation gives rise to the "looking-for" which then directs our attention. The degree of resonance among these activities corresponds to the vividness of the consciousness. The quality of experience as "stretched out" over the three temporal moments—present, immediately retained just-past, and anticipated future—also leads to a feeling of an inseparable blending of feeling with the intending of an object, and even a fuzziness in distinguishing subject from object (as when we attribute the phenomenal redness of an object to the object itself, as if the red were pasted to the surface of the object, or when we attribute the mood that an object produces in us to the object itself). This blending accounts for a good deal of the fact that we usually focus our attention on the object of experience rather than our subjective experiencing of it. Here again, what is subjective and what is objective about our experience of reality is in principle not a clear or sharp distinction.

Consciousness cannot be understood as a passive causal result of the actions of the substratum elements that make up the components forming the higher-order process that allows for the active nature of consciousness. Consciousness is self-organizational in a biological way. The higher-level process seeks out, appropriates, re-arranges, and within certain limits even replaces the micro-level constituents needed to enact its dynamical patterns of organization. Consciousness is not simply caused by its micro-level physical components, but neither is it separable from them. The enactive approach meets this demand by conceiving of consciousness as a self-organizing process that is inseparable from its micro-constituents in the sense that it could not occur without them, yet it is not passively caused by the actions of those substrata. But this theory in turn requires a careful scientific grounding of the causal relations required to maintain the regularly observed "causal closure" at the micro level, as our next section will discuss.

The compatibility of dynamical systems with micro-level mechanistic explanation

There is now a mass of empirical evidence pointing to the thesis that processing occurs in a *conscious* way only when it is very globally distributed in the brain. For example, we know that, when impulses caused by optic stimulation set up patterns of activity in the occipital lobe, but without coordinated limbic and frontal-cortex activity, no perceptual consciousness results from the occipital activity (Posner 1990; Damasio et al. 2000; Eslinger and Damasio 1985; Nauta 1971; Luria 1980). Similarly, the transition from sleep to waking consciousness requires that the activities of the hypothalamus and cortex achieve a pattern of synchronization or coordination which was not present during sleep (Asimov 1965: 193; Ellis 1986: 46-52). When we are conscious of dream images during sleep, both efferent and afferent activity throughout the brain are detected, whereas during non-dreaming sleep both the afferent activity and some of the efferent activity are comparatively much less pronounced (Winson 1986: 46ff; Restak 1984: 315-333; Richardson 1991; Jouvet 1967).

Another example is provided by the 1/3-second time delay from the activation of the occipital lobe (in response to a novel stimulus) to the presence of perceptual consciousness of the object, the latter correlating with coordinated limbic, frontal, parietal, and occipital activity (Aurell 1989; Runeson 1974: 14; Srebro 1985: 233-46). EEG and other electrical measures show that parietal activation does not occur until 1/3 second after the occipital activity, which by itself does not produce consciousness of the object. As we saw earlier, the activation does not merely “travel” from the occipital to the adjacent parietal area; if it did, the distance involved would be traversed in much less time than 1/3 second. Instead, before perceptual consciousness can occur, the limbic system must be aroused, and it in turn must activate the frontal lobe to begin formulating questions about what the nature of the interesting or important environmental stimulus might be, which then activates ideas and/or images in the form of anticipations of possible perceptions with the help of the parietal lobe (Ellis 1995, Ch. 1; Luria 1980; Posner and Rothbart 1992). Only as a result of this symphonic orchestration of global activity can the activity of the parietal lobe be matched against what is happening in the occipital lobe to see whether the image or idea hypothesized is actually instantiated in the environment. If so, perceptual consciousness of the corresponding object occurs. If not, a mere mental image of the object experienced as non-present occurs in consciousness. In either event, consciousness occurs only when brain activities are globally coordinated. What these examples and many others suggest is that consciousness requires globally distributed processes in the brain, combining local mechanisms which under different circumstances would be active in various non-conscious processes. Luria (1980), Posner (1990), Posner and Rothbart (1992), and many other neurologists are now convinced that the prefrontal lobe plays a crucial role in the process of directing attention to what is important. What makes it do so is that it receives rich input from the limbic system (importantly involved in motivational feelings and other “valuations”), and then sends signals that coordinate the remainder of the cortex to be consciously aware of the arousing situation and to devise ways to deal with it.

Neither dualism, nor psychophysical identity theory, nor epiphenomenalism works as an explanation of the relation between consciousness and its physiological correlates, because the modernist concept of atomistic-reductionism does not allow a process to affect the behavior of its own substratum elements, but requires that a process must be caused by the interaction of the discrete movements of its substratum elements, each of which has a sufficient causal explanation of its own, so that the pattern of consciousness, paradoxically, can have no causal power.

But the enactive approach avoids this aspect of the problem of causal closure, because it does allow that a *process* can have causal power. The necessary and sufficient causal antecedent of an outcome in a self-organizing dynamical system is a previous *pattern* of functioning, which can be realized by alternative sets of micro-constituents. None of the particular micro-constituents are strictly necessary, because others could have been used to maintain the same process. But not only is the process multiply realizable. In self-organizing systems, the entity defined by the overall process *makes use* of the fact of multiple realizability by seeking out, appropriating, and replacing the micro-constituents needed to keep the process going in its definitive pattern.

Causal closure is nonetheless still maintained in this case because a process is inseparable, *under the given circumstances* (but only under those given circumstances), from the behavior of its actual micro-constituents or substratum elements. So the process will also be necessary and sufficient for whatever its substratum elements are necessary and sufficient for, under those specific circumstances. Yet this does not necessarily imply that the process is caused by its substratum, or that it is identical with it. Many things are true of a process which are not true of its substratum elements, even taken collectively. For example, a wave on the ocean may travel many miles in a horizontal direction, while its substratum elements, the movements of particles of water, are very small vertical oscillations.

The reason the process can have this kind of power to rearrange its own substratum elements in such a "plastic" way is that causal antecedents are necessary and sufficient for their consequents *only when certain background conditions are in place.* The power of the process is that it can alter the background conditions themselves, because the process is arranged in such a way that shunt mechanisms will kick in when a preferred causal relation (such as a chemical reaction) fails to occur. There are multiple cause-and-effect sequences that can lead to the same pattern being maintained, depending on which sets of background conditions are in place. What is really necessary and sufficient for a given outcome is that a previous *process* have been in place—not simply a series of micro-reactions, as if those micro-reactions were not dependent on a larger set of background conditions being in place.

The process-substratum relation in the case of consciousness is different from the relationship between a wave and the physical medium through which the wave passes, such as a sound wave, in one crucial respect. Consciousness, unlike a sound wave or a wave in the ocean, is a purpose-directed process. Merleau-Ponty (1942/1967: 47ff) defines a purposeful activity as one in which the organism's overall pattern of activity acts in such a way as to rearrange and readjust its various parts in order to maintain or enhance the overall pattern. Purely mechanical processes do not seem to behave in this way. A thermostat, while it will adjust its overall pattern to feedback from the environment, does not seem to be a purpose-directed system because, when one of its parts ceases to function or is removed, the thermostat does not act in such a way as to replace the missing part or try to compensate for its absence; it simply quits functioning. The thermostat does not "care," in this non-conscious sense of "care," whether it achieves its ultimate objective or not. It functions or not purely as an additive juxtaposition of the functioning of its parts.

It becomes increasingly clear, as we study the brain, the ecosystem, and the concept of “living organisms” in biology, that at least many patterns of activity maintain their organizational structure across replacements of their own substrata. As Merleau-Ponty suggests, an organism will often rearrange the overall configuration of its parts if an imbalance is created in one part which disrupts the functioning of the whole. "'Forms' . . . are defined as total processes whose properties are not the sum of those which the isolated parts would possess. . . . We will say that there is form whenever the properties of a system are modified by every change brought about in a single one of its parts and, on the contrary, are conserved when they all change while maintaining the same relationship among themselves" (Merleau-Ponty 1942/1967: 47).

Examples of self-directed neurophysiological reorganization following localized brain injury or trauma can be found in Restak (1984: 360ff). Kandel and Schwartz (1981) also emphasize the “plasticity” of the brain in reorganizing itself to accomplish its objectives by getting around disruptions in one way or another. For example, if brain cells of an embryo are transplanted to a different region of another embryo, they are transformed into cells appropriate to that region. This plasticity in the realizability of the mental functions of living beings has been discussed by Bechtel and Mundale (1999), Gillett (2003), Horgan (1992) and Bickle (2006). An example of multiple realization is that the organism's desires intend to remove the inevitability of electrochemical imbalance within the organism, not merely by eliminating this or that electrical imbalance (for example, in cases where to do so would only transmit the imbalance from one part of the nervous system to another, or from one bodily system to another), but rather by changing the context which renders the imbalance inevitable—for example, by spatially removing the entire organism from the disturbing stimulus, by destroying the disturbing stimulus, or by finding or creating a stimulus in relation to which the whole organism's balance can be restored.

Conclusion

The twentieth century philosophy of mind made every effort to remain tenaciously bottom-up. Cognitive functions have been explained as “responses” to incoming “stimuli,” with the stimuli combining in complex ways to mechanically cause the response. The response is thus a purely passive change, brought about by the stimulus. As in the characteristic twentieth century approach to natural science, here too the only inertia is an inertia of passivity; nothing would move or change unless acted upon by an outside force.

In order to overcome the problems we have just outlined, an adequate conception of consciousness must reopen these questions with regard to ontology and the theory of causation. We must develop a theory in which purposeful processes are able to appropriate their needed substratum elements, rather than merely being passive epiphenomena of them or ontologically identical with them. This in turn will require the development of a workable account of how it is that certain activities can be "purposeful" in a scientifically intelligible universe. The mainstream of twentieth century psychology predominantly turned its back on this problem. Purposeful activity was explained away as a purely mechanical process that only *appears*, anthropomorphically, as if it were purposeful. The standard explanation was that we view a mechanical process as if it were purposeful because we view it as if it were conscious, like ourselves, and we imagine that if we were to engage in that activity, we would be doing so with the consciousness of some purpose in mind. But to characterize a process as purposeful is not to anthropomorphize. The human organism was purposeful before it was conscious—for example, during embryonic development. Consciousness is not necessary to purposefulness, even in the human organism. So purposefulness cannot be explained simply as the addition of consciousness to a process that otherwise could be explained simply as one that displays certain "tendencies" to accomplish certain results, as if the only difference between a purposeful and a non-purposeful process were that, in the latter, there is conscious awareness of an underlying sequence of purely passive reactions. The new theories of self-organization that are now available can be applied to understanding the causal power of a process over its own components, and thus the teleological nature of consciousness as well as many other purpose-directed processes in nature, such as the migration of birds, or the balance of ecosystems.

References

Asimov, Isaac. 1965. *The Human Brain.* New York: Mentor.

Aurell, Carl G. 1989. Man's triune conscious mind. *Perceptual and Motor Skills* 68: 747-754.

Bachmann, Talis. 2000. *The Microgenetic Approach to the Conscious Mind*. Amsterdam: John Benjamins.

Bechtel, William, and Jennifer Mundale. 1999. Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science* 66: 175-207.

Bickle, John. 2006. Multiple realizability. *The Encyclopedia of Cognitive Science.* New York: Wiley.

Bickle, John. 2008. *Psychophysical Reduction: The New Wave.* Boston: MIT/Bradford.

Chalmers, David. 1995. Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2: 5-22.

Corbetta, M., F.M. Meizen, S. Dobmeyer, G.L. Schulman, and S.E. Petersen. 1990. Selective attention modulates neural processing of shape, color and velocity in humans. *Science* 248: 1556-1559.

Cytowic, Richard. 1993/1995. *The Man Who Tasted Shapes.* New York: Warner.

Damasio, Antonio. 1994. *Descartes' Error.* New York: Putnam.

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J. 2000. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience 3*: 1049-1056.

Dascal, Marcelo. 1987. Language and reasoning: Sorting out sociopragmatic and psychopragmatic factors. J.C. Boudreaux, B. W. Hamill, and R. Jernigan, eds. *The Role of Language in Problem Solving 2*. Elsevier: North-Holland, 183-197.

Dennett, Daniel. 1996. *Kinds of Minds.* New York: Basic Books.

Ellis, Ralph D. 1986. *An Ontology of Consciousness*. Dordrecht: Springer/Kluwer/Martinus Nijhoff.

Ellis, Ralph D. 1995. *Questioning Consciousness: The Interplay of Imagery, Cognition and Emotion in the Human Brain.* Amsterdam: John Benjamins.

Ellis, Ralph D. 2005. *Curious Emotions.* Amsterdam: John Benjamins.

Ellis, Ralph D., and Natika Newton. 1998. Three paradoxes of phenomenal consciousness. *Journal of Consciousness Studies* 5: 419-42.

Ellis, Ralph D., and Natika Newton. 2010. *How the Mind Uses the Brain.* Chicago: Open Court.

Eslinger, Paul, and A.R. Damasio. 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation. *Neurology* 35: 1731-1741.

Farah, Martha. 1989. The neural basis of mental imagery. *Trends in Neuroscience* 12: 395-399.

Freeman, Walter. 2001. *How Brains Make Up Their Minds.* New York: Columbia University Press.

Gendlin, Eugene. 1992. Thinking beyond patterns. B. den Ouden and M. Moen, eds. *The Presence of Feeling in Thought.* New York: Peter Lang.

Gibson, J.J. 1986. *An Ecological Approach to Visual Perception.* Hillsdale, N.J.: Lawrence Erlbaum.

Gillett, Carl. 2003. The metaphysics of realization, multiple realizability, and the special sciences. *The Journal of Philosophy* 100: 591-603.

Goldstein, Irwin. 1994. Identifying mental states: a celebrated hypothesis refuted. *Australasian Journal of Philosophy* 72: 46-62.

Globus, Gordon. 1992. Towards a noncomputational cognitive neuroscience. *Journal of Cognitive Neuroscience 4:* 299-310.

Globus, Gordon. 2003. *Quantum Closures and Disclosures*. Amsterdam: John Benjamins.

Haines, D., E., Dietrich, E., Mihailoff, G.A., and McDonald, E. F. 1997. Cerebellar-hypothalamic axis: Basic circuits and clinical observations. J. Schmahmann, ed. *The Cerebellum and Cognition* (pp. 84-110). New York: Academic Press.

Hanze, Martin and Friedrich Hesse. 1993. Emotional influences on semantic priming. *Cognition and Emotion* 7: 195-205.

Helmholtz, Hermann. 1962. *Helmholtz's Treatise on Physiological Optics,* J.P.C. Southall (trans). New York: Dover.

Hull, C. L. 1952. *A Behavior System.* New Haven: Yale University Press.

Ito, Masao. 1993. Movement and thought: Identical control mechanisms by the cerebellum. *Trends in the Neurosciences* 16: 448-450.

Jackendoff, Ray. 1987. *Consciousness and the Computational Mind.* Cambridge: MIT Press.

Jeannerod, M. 1994. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences* 17: 187-244.

Jeannerod, M. 1997. *The cognitive neuroscience of action.* Oxford: Blackwell.

Jibu, Mari, and Kunio Yasue. 1995. *Quantum Brain Dynamics and Consciousness*. Amsterdam: John Benjamins.

Johnson-Laird, Philip N., and R.M.J. Byrne. 1991. *Deduction.* Hillsdale, N.J.: Erlbaum.

Jouvet, Michael. 1967. Neurophysiology of the states of sleep. *Physiological Review* 47: 117-127.

Kandel, Eric, and James Schwartz. 1981. *Principles of Neural Science.* New York: Elsevier-North Holland.

Kauffman, Stuart. 1993. *The Origins of Order.* Oxford: Oxford University Press.

Kelso, Scott. 1995. *Dynamic Patterns: The Self-organization of Brain and Behavior*. Boston: MIT/Bradford.

Kim, J. 1992. Multiple realization and the metaphysics of reduction. *Philosophy and Phenomenological Research 52:* 1-26.

Kim, J. 1993. The nonreductivist's troubles with mental causation. J. Heil and A. Mele, eds. *Mental causation* (pp. 189-210). Oxford: Oxford University Press.

Kim, J. 1998. *Mind in a Physical World: an Essay on the Mind-body Problem and Mental Causation*. Cambridge, Mass.: MIT Press.

Lavy, Edith and Marcel van den Hout. 1994. Cognitive avoidance and attentional bias: Causal relationships. *Cognitive Theory and Research* 18: 179-194.

LeDoux, Joseph. 1996. *The Emotional Brain.* New York: Simon & Schuster.

Legrenzi, P., V. Girotto, and P.N. Johnson-Laird. 1993. Focussing in reasoning and decision making. *Cognition* 49: 37-66.

Lethin, Anton. 2002. How do we embody intentionality? *Journal of Consciousness Studies* 9: 36-44.

Lethin, Anton. 2005. Covert agency with proprioceptive feedback. *Journal of Consciousness Studies* 12:96-114.

Levine, Joseph. 1983. Materialism and qualia; the explanatory gap. *Pacific Philosophical Quarterly* 64: 354-61.

Logan, G.D. 1980. Attention and automaticity in stroop and priming tasks: Theory and data. *Cognitive Psychology* 12: 523-553.

Luria, Alexander R. 1980. *Higher Cortical Functions in Man,* 2nd ed. New York: Basic Books.

Mack, A., and Rock, I. 1998. *Inattentional Blindness.* Cambridge: MIT/Bradford.

McHugh, D.E. and A.T. Bahill. 1985. Learning to track predictable target waveforms without a time delay. *Investigative Ophthalmology and Visual Science* 26: 932-937.

Merleau-Ponty, Maurice. 1942/1967. *The Structure of Behavior.* A. Fischer (trans). Boston: Beacon.

Merleau-Ponty, Maurice. 1941/1962. *Phenomenology of Perception.* Colin Smith (trans). New York: Humanities Press.

Monod, Jacques. 1971. *Chance and Necessity.* New York: Random House.

Moore, G.E. 1900/1956. *Principia Ethica.* Cambridge: Cambridge University Press.

Natsoulas, Thomas. 1994. What is wrong with appendage theory of consciousness. *Philosophical Psychology* 6: 137-154.

Nauta, Walle. 1971. The problem of the frontal: A reinterpretation. *Journal of Psychiatric Research* 8:167-187.

Newton, Natika. 1993. The sensorimotor theory of cognition. *Pragmatics and Cognition* 1: 267-305.

Newton, Natika. 1996. *Foundations of Understanding.* Amsterdam: John Benjamins.

Newton, Natika. 1982. Experience and imagery. *Southern Journal of Philosophy* 20: 475-487.

Newton, Natika. 1989. Visualizing is imagining seeing: A reply to White. *Analysis* 49: 77-81.

Newton, Natika. 1991. Consciousness, qualia, and reentrant signalling. *Behavior and Philosophy* 19: 21-41.

Newton, Natika. 1992. Dennett on intrinsic intentionality. *Analysis* 52: 18-23.

Newton, Natika. 2000. Emotion in a dynamical system. R. Ellis and N. Newton, eds. *The Caldron of Consciousness.* Amsterdam: John Benjamins.

Noë. Alva. 2006. *Action in Perception.* Boston: MIT/Bradford.

Panksepp, Jaak. 1998. *Affective Neuroscience.* Oxford: Oxford University Press.

Panksepp, Jaak. 2012. defense of multiple core affects. P. Zachar and R. Ellis,, eds. *Categorical and Dimensional Models of Affect.* Amsterdam: John Benjamins.

Pardo, J.V, P.J. Pardo, K.W. Janer, and M.E. Raichle. 1990. The anterior cingulate cortex mediates processing selection in the stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences* 87: 256-259.

Posner, Michael I. 1980. Orienting of attention. *Quarterly Journal of Experimental Psychology* 32: 3-25.

Posner, Michael I. 1990. Hierarchical distributed networks in the neuropsychology of selective attention. A. Caramazza (ed). *Cognitive Neuropsychology and Neurolinguistics: Advances in Models of Cognitive Function and Impairment.* New York: Plenum, 187-210.

Posner, Michael I., and S.E. Petersen. 1990. The attention system of the human brain. *Annual Reiew of Neuroscience* 13: 25-42.

Posner, Michael I. and Mary K. Rothbart. 1992. Attentional mechanisms and conscious experience. A.D. Milner and M.D. Rugg, eds. *The Neuropsychology of Consciousness.* London: Academic Press.

Restak, Richard. 1984. *The Brain.* New York: Bantam.

Rhodes, Gillian and Tanya Tremewan. 1993. The Simon then Garfunkel effect: Semantic priming, sensitivity, and the modularity of face recognition. *Cognitive Psychology* 25: 147-87.

Richardson, John. 1991. Imagery and the brain. C. Cornaldi and M. McDaniels, eds. *Imagery and Cognition.* New York: Springer, 1-46.

Rolls, E. 1999. *The Brain and Emotion.* Oxford: Oxford University Press.

Runeson, Sverker. 1974. Constant velocity—not perceived as such. *Psychological Research* 37: 3-23.

Schiffman, Harvey Richard. 2001. *Sensation and Perception: An Integrated Approach.* New York: John Wiley.

Smart, J.J.C. 1970. Sensations and brain processes. *Philosophical Review 68,* Reprinted in C.V. Borst, ed. *The mind-brain identity theory* (pp. 141-56). London: Macmillan.

Spence, K. W. 1956. *Behavior Theory and Conditioning.* New Haven: Yale University Press.

Spence, S.A., and Frith, C.D. 1999. Towards a functional anatomy of volition. *Journal of Consciousness Studies 6:* 11-29.

Springer, Sally, and Georg Deutch. 1989. *Left Brain, Right Brain.* New York: W.H. Freeman.

Srebro, Richard. 1985. Localization of visually evoked cortical activity in humans. *Journal of Physiology* 360: 233-246.

Thelen, E., and Smith, L. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Massachusetts: MIT Bradford.

Thompson, Evan. 2007. *Mind in Life.* New York: Belknap.

Valenstein, Elliot. 2002. *Blaming the Brain.* Chicago: Free Press.

Varela, Francisco, Evan Thompson, and Eleanor Rosch. 1991/1993. *The Embodied Mind.* Cambridge: The MIT Press.

Winson, Jonathan. 1986. *Brain and Psyche.* New York: Random House.

Wrightson, Philip. 1989. Management of disability and rehabilitation services after mild head injury. In H.S. Levin, H.M. Eisenberg, and A.L. Benton, eds. *Mild Head Injury*: 245-256. Oxford: Oxford University Press.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0009

Spatial Content and Motoric Significance

Robert Briscoe
Department of Philosophy
Ohio University
rbriscoe[]gmail.com

Received December 2013; accepted September 2014; published Autumn 2014.

Abstract

According to "actionism" (Noë 2010), perception constitutively depends on implicit knowledge of the way sensory stimulations vary as a consequence of the perceiver's self-movement. My aim in this contribution is to develop an alternative conception of the role of action in perception present in the work of Gareth Evans using resources provided by Ruth Millikan's biosemantic theory of mental representation.

Keywords: enactivism; egocentric spatial representation; visuomotor control; biosemantics; Gareth Evans.

1. Introduction

The sensorimotor contingency theory of perception (O'Regan & Noë 2001, Noë 2004, Noë 2010, O'Regan 2012) or "actionism," as it had been more recently called, is central to numerous enactivist projects in the philosophy of mind and cognitive science.[42] According to actionism, perceptual experience constitutively depends on knowledge of the way sensory stimulation varies as a result of self-initiated, bodily movement. "For something to be visible," Alva Noë writes, "…is for it to show up as standing to us in a relation of, as I will put it, sensorimotor perturbability. If you see something, then movements of your eyes or body will affect the way you experience it" (2010: 248). Actionism it is important to emphasize, however, distances itself from the idea that vision is *for* action, that vision is functionally dedicated, in whole or in part, to the guidance of spatially directed bodily movement: "actionism does not claim that visual awareness depends on visuomotor skill, if by 'visuomotor skill' one

[42] Enactivism isn't so much a single, unified research perspective as it is a cluster of more or less closely related explanatory frameworks and philosophical methodologies, including but not limited to noncomputational and nonrepresentational dynamical systems theory, embodied cognitive science, ecological psychology, vehicle externalism, naïve realism, and certain forms of existential phenomenology.

means the ability to make use of vision to reach out and manipulate or grasp. Our claim is that seeing depends on an appreciation of the sensory effects of movement (not, as it were, on the practical significance of sensation).... Actionism is not committed to the general claim that seeing is a matter of knowing how to act in respect of or in relation to the things we see" (Noë 2010: 249).[43] To experience the three-dimensional shape, size, or orientation of a rock on the beach, on this view, it thus isn't necessary to understand, in a practical sense of understanding, how you would need to move your body if it were your purpose, say, to approach the rock, or to reach for it, or to pick it up in a certain way, or even to look in its direction. What is necessary, rather, is to understand how retinal stimulations caused by the light reflected from rock would change were you to act in any of these or other ways.[44] For the

---

[43] There are at least two distinct senses in which seeing could be *for* action. Seeing could be for action in the sense that one of its psychological-kind-individuating functions is to guide visuomotor actions, e.g., by supplying motor programming system with information about the spatial properties of visible objects. Alternatively (and much less plausibly), vision could be for action in the sense that action guidance is its *only* function. Actionism denies that seeing is for action in either sense. I am grateful to any anoymous referee for prompting me to make this point.

[44] One source of empirical evidence for this view comes from studies of visuomotor development in the absence of normal, reafferent visual stimulation. Held & Hein 1963 performed a now famous experiment in which pairs of kittens were harnessed to a carousel in a cylindrical chamber. One of the kittens in each pair was harnessed in such a way that it was able to engage in free circumambulation. The other kitten was suspended in the air in a metal gondola. When the first kitten walked, both kittens moved in a circle and received identical visual stimulation. However, only the first kitten received reafferent visual feedback as the result of self-movement. Held and Hein reported that only *mobile* kittens developed normal depth perception—as evidenced by their unwillingness to step over the edge of a visual cliff. Noë (2004) argues that this experiment supports the enactive approach: in order to develop normal visual depth perception, cats and other animals have to learn the sensory consequences of their own movements.

There are reasons to be skeptical of this assessment. For one thing, there is evidence that passive transport in the gondola may have disrupted the development of the kittens' innate paw placing responses to visually perceived surfaces (Ganz 1975: 206). Second, the fact that passive kittens were prepared to walk over the edge of a visual cliff doesn't show that their visual experience of depth was abnormal. Rather, as Jesse Prinz (2005) points out, it may only indicate that they "did not have enough experience walking on edges to anticipate the bodily affordances of the visual world." Evidence for this interpretation comes from studies of visual space perception in human infants. Campos *et al.* 1992 found that infants, ranging from 6.5 to 8.5 months, exhibit wariness of heights when lowered onto the deep side of a visual cliff only when they had prior experiences of crawling or using a walker. There is no empirical reason to think, however, that pre-locomotor infants in this age cohort are unable to perceive three-dimensional spatial layout (Campos 2000). Indeed, the available evidence suggests that just the opposite is the case (for a review, see Kellman & Arterberry 2006). By one month of life, infants will blink defensively when presented with optical expansion patterns that normally signify an approaching object (Nanez & Yonas 1994); by the end of the fourth month, binocular disparity is operative (Braddick & Atkinson 1983); and by the seventh month, infants, regardless of previous locomotor experience, are sensitive to the "pictorial" cues of occlusion, familiar size, and height in the visual field, as evidenced by their visually guided reaching behavior (Granrud & Yonas 1984, Granrud *et al.* 1985, Arterberry 2008). Summarizing decades of developmental evidence, Kellman and Arterberry write: "Before crawactionist, the possession of visuomotor skill or "know-how" isn't required for visual awareness of the way objects fill out surrounding space.

I have elsewhere engaged in detailed criticism of the actionist approach (Briscoe 2008a, Briscoe *forthcoming*, Briscoe and Grush *in preparation*). Like a number of other philosophers, I am skeptical both about its internal coherence and empirical tenability (Block 2003, 2005, 2012; Prinz 2006, 2012; Matthen 2006; Schwitzgebel 2006; Clark 2009, 2012). I won't attempt to summarize the state of the debate here. My aim, instead, is to lay out an alternative conception of the role of action in perception present in the work of Gareth Evans (1982, 1985), one that I think is far more consistent with mainstream empirical research in perceptual psychology and cognitive neuroscience and that has yet to be developed, I think, in a sufficiently careful and systematic way. Unlike proponents of actionism, Evans does not look to knowledge of the proximal sensory consequences of movement in order to explain how perception acquires its spatial content. Rather, he looks to the functional role played by perception in adapting the agent's bodily actions to the spatial layout of the distal environment. Unlike actionism, Evans' theory *is* committed to the claim that "seeing is a matter of knowing how to act in respect of or in relation to the things we see" and, so, to the constitutive dependence of visuospatial awareness on the possession of visuomotor skill.

Here is an overview of the rest of this paper. In section 2, I present an interpretation of four claims central to Evans's theory of the egocentric spatial content of perception. I also answer objections to Evans's theory that arise from a failure to distinguish between the objective spatial content of a perceptual experience and the experience's motoric significance for the perceiving subject. In section 3, I then show that Evans' theory can be helpfully elaborated using resources provided by Ruth Millikan's biosemantic theory of mental representation.

2. Evans on the role of action in perception

The idea that action and perception are closely related is neither new, nor especially radical. Many philosophical and psychological theories of vision in the last 300 years have looked to capacities for embodied, visuomotor action to explain how visual experience acquires its spatial representational content (for a review, see Briscoe and Grush *in preparation*). One historically important source of motivation for this approach is the empiricist doctrine that

ling, infants perceive depth; what may change with crawling experience is the coordination of depth and surface perception with their own motion in space" (1998: 261). An analogous conclusion can plausibly be drawn with respect to Held and Hein's kittens.

vision must be "educated" by the sense of touch—understood as including both kinaesthesis and proprioceptive position sense—if the former is to acquire its outward, three-dimensional spatial significance. Normal vision, Berkeley influentially argued in his *New Theory* (1709/2008), results only when tangible ideas of distance (derived from experiences of unimpeded movement) and solid shape (derived from experiences of contact and differential resistance) are elicited by the visible ideas of light and color with which they have been habitually connected. A long line of philosophers including Condillac, Reid, Smith, Mill, Bain, and Dewey accepted the basics of Berkeley's account of the relation between sight and touch.

A second important source of motivation for action-oriented approaches to perception is teleological. From an biological or evolutionary standpoint, it is reasonable to think that vision is *for* action, that its preeminent biological function is to adapt an animal's bodily movements to the properties of the environment that it inhabits. This view is widely accepted in the neuroscience of visuomotor control: "the functional organization of the visual system (like the rest of the brain)," Melvyn Goodale writes, "has been ultimately shaped by the role it plays in the control of movement" (Goodale 2011: 1568). It is also clearly reflected in externalist or "anti-individualist" approaches to perceptual content in the philosophy of mind. "The representational content of an animal's perceptual states," as Tyler Burge puts it, "is individuated partly in terms of what causes those states and how those states enable the animal to cope with specific types entities in its environment. Successful interactions help ground individuation of perceptual states partly in terms of representational content" (2005: 5). Seeing, of course, subserves a variety of other important purposes besides the guidance of actions—its contents are both directive and descriptive (Millikan 2004)—but this is plausibly one of its biologically central and psychological-kind-individuating functions.

Yet a third source motivation comes from skepticism about the explanatory adequacy of alternative approaches. Most philosophers of mind, it is fair to say, now concur that representation is a *functional* kind. A mental state is a vehicle of representation content only if it used in certain ways, only if it plays a certain role in the agent's cognitive economy. It seems unlikely, however, that perceptual states acquire spatial contentfulness in virtue of their role in propositional inference (for instance, in propositional, spatial reasoning). Indeed, capacities for propositional inference seem neither necessary nor sufficient for perceptual spatial representation. They don't seem necessary because young human infants and many animals that lack capacities for propositional inference are evidently three-dimensional space perceivers. And they don't seem sufficient because there are good reasons to think that the spatial content of perception is, in general, nonconceptual—and, so, constitutively independent of having capacities for propositionally articulated reasoning (for discussion, see the essays collected in Gunther 2003, Bermúdez 2007, and Burge 2010, chap. 11).

In Chapter 6 of *The Varieties of Reference* (1982), Gareth Evans argues that perceptual states acquire their nonconceptual spatial contents in virtue of the role they play in the guidance of spatially directed motor actions. The approach Evans develops is structured by four main claims:

> (1) The spatial content of perceptual experience is subject-relative or "egocentric" (if not exclusively so). "The subject hears the sound as coming from such-and-such a position, but how is the position to be specified? Presumably in *egocentric* terms.... These terms specify the position of the sound in relation to the observer's own body" (1982: 155).
>
> (2) We perceive egocentric spatial properties using the same spatial coding system or "frame of reference" as is used for purposes of forming and implementing our intentions for object-directed bodily actions. "Egocentric spatial terms are the terms in which the content of our spatial experiences would be formulated, and those in which our immediate behavioral plans would be expressed" (1982: 154).
>
> (3) Our perception of an object's egocentric spatial properties is constitutively connected with having certain *dispositions* to engage in bodily actions targeted on or otherwise directed in relation to the object. "[W]e must say that having spatially significant perceptual information consists at least partly in being disposed to do various things" (1982: 155).
>
> (4) Perceptual information about an object's position in egocentric space is not information about a special kind of space, but rather *information* of a special kind about space. "It is perfectly consistent with the *sense* I have assigned to [egocentric spatial] vocabulary that its terms should *refer* to points in a public three-dimensional space" (1982: 157).

My aim in this section is to interpret these claims by clarifying them and relating them to one another. I shall also answer objections to Evans's account that arise, I suggest, from a failure to distinguish between the objective, egocentric spatial content of a visual experience and its motoric significance for the perceiving subject (Claim 4). Although I shall be focusing on the case of conscious vision, points made here are intended to generalize to other sensory modalities.

Let's begin with Claim 1. According to Claim 1, visual experience represents the egocentric spatial properties of visible objects and surfaces, that is, their spatial relations to the perceiving subject. Thus, when you see a plate on the table, you see among other things its direction and distance from *your* own body as well as its three-dimensional orientation relative to *your* line of sight. Claim 1, it is important to stress, does not exclude the possibility that visual experiences also represent various allocentric spatial properties and relations.

E.g., in addition to seeing the plate's location relative to your own body, you may also see its position relative to a nearby fork and saltcellar.

Three remarks should be made about Claim 1. First, there is psychophysical evidence that at close range and under ecologically normal viewing conditions our ability consciously to perceive egocentric spatial layout is highly accurate. While perceived distances are significantly foreshortened for objects located more than 30 meters away, in what Cutting and Vishton (1995) call *vista space*, the egocentric distances of objects located up to 2 meters away, in *personal space*, are perceived with nearly metric accuracy.[45] Although proponents of the dual systems model of visual processing (Milner & Goodale 1995/2006, Goodale & Milner 2004) have maintained that conscious vision does not make use of an egocentric spatial coding system, this claim, I would suggest, is flatly inconsistent with mainstream psychophysical work in perceptual psychology (see Briscoe 2008b, 2009 and Briscoe & Schwenkler *forthcoming*).

Second, Claim 1 is a claim about the *spatial contents* of visual experience—about which objective (but subject-relative) spatial properties are represented in visual experience—and as such does not by itself entrain any specific commitments about the system of egocentric spatial representation used to encode those contents. Indeed, possible perceptual mental representations with a given egocentric spatial content, as Christopher Peacocke points out, stand in a *many-one* relation to the content itself (1992: 65). Knowing *which* spatial properties are encoded by a system of spatial representation *S* does not by itself tell us *how* those properties are encoded by *S*.

Third, Claim 1 by itself does not identify the location of the self or *ego* in relation to which egocentric spatial properties are supposed to be represented in visual experience. Indeed, it does not indicate whether there is a single, privileged locus in (or on) the body that counts as the center of visual egocentric space and, so, does not indicate whether the ego *qua* perceiving subject is literally a *point* of view—in the world only "geometrically," as John McDowell puts it (1994, 104). This means that, in addition to the leaving the nature of the spatial coding system used in visual experience indeterminate, Claim 1 also leaves indeterminate how that system is to be aligned with the perceiving subject's body.

Although this may seem like a straightforward phenomenological question, it is not. In fact, different philosophers influenced by Evans's account have given quite different answers to it. Christopher Peacocke (1992), for instance,

[45] The difference in precision is a function of the spatial information available to the visual system. Estimates of depth for objects in personal space are powerfully constrained by stereopsis, convergence, and accommodation. As distances lengthen, these binocular sources of depth information drop off in effectiveness, and the visual system must rely increasingly on somewhat less precise monocular or "pictorial" cues in the light sampled by the eyes.

maintains that visual experiences represent the way visible surfaces are arrayed in three-dimensional space relative to a point placed in the subject's torso. By contrast, Jose Bermúdez (1998: 2005) identifies the *ego* at the center of visual egocentric space with the apex of the solid angle of the visual field, while Quassim Cassam (1997) identifies the *ego* with the living, acting body as a whole (what he calls the "bodily self").

One reason why visual phenomenology is not decisive here, I would suggest, is that there is disagreement about what counts as distinctively *visual* experience of space. In particular, when we say that we visually perceive the spatial relations in which certain objects stand to ourselves, what role does proprioception play in the representation of those spatial relations? Consider the case of seeing an object's (radial) direction. If Bermúdez is right, then visual experience only represents the object's direction from you relative to a point midway between your eyes. In this case, your representation of the tree's egocentric direction is fully independent of concurrent proprioceptive information, whether conscious or nonconscious, about the spatial configuration of the rest of your body.

Peacocke, by contrast, maintains that the directional axes used to specify the spatial content of a visual experience originate from a point in the center of the torso. "The appropriate set of labeled axes," he writes, "captures distinctions in the phenomenology of experience itself. Looking straight ahead at Buckingham Palace is one experience. It is another to look at the palace with one's face still toward it but with one's body turned toward a point on the right. In this second case the palace is experienced as being off to one side from the direction of straight ahead, even if the view remains exactly the same as in the first case" (1992: 62). The visual system, however, initially encodes an object's location relative to the eye (that is, in retinocentric coordinates). Representing an object's location in visual experience relative to the torso thus presupposes prior integration of visual information about the object's eye-relative location with proprioceptive information about the spatial configuration of the body, in particular, information about the direction of gaze and the orientation of the head. We could call this a "visuo-proprioceptive" representation of the palace's direction.

Something similar, of course, seems possible in respect of other proprioceptively represented parts of the body, for example, the head, shoulder, or hand. If so, then there needn't be a single bodily locus that, in general, counts as *you* for purposes of characterizing your visually perceived spatial relations to an object. The *ego* at the center of perceptual egocentric space, as suggested by Cassam, may spread to encompass the body as a whole.[46]

[46] This does not assume that the perceiver is delivered in visual experience with a complete and uniformly detailed representation of an object's location relative to *every* part of her body at the

I will not attempt to adjudicate between these different views here.[47] It is enough to point out that which egocentric spatial relations one takes to be represented in visual experience depends on where one locates the *ego* at the center of visual egocentric space, and this, in turn, depends in part on how one conceives of the relationship between conscious vision and bodily proprioception.

Claim 1 was a claim at the level of objective spatial representational content. Claim 2, by contrast, is a claim at the level of spatial representational format. According to Claim 2, we perceive egocentric spatial properties using the same spatial coding scheme as is used for purposes of forming and implementing our intentions for spatially directed movement and action. As Evans puts it, "Egocentric spatial terms are the terms in which the content of our spatial experiences would be formulated, and those in which our immediate behavioral plans would be expressed" (1982: 154).

One motivation for Claim 2 is presumably that, in order to form intentions for bodily actions directed in relation to an object, it is necessary to locate the object's position in three-dimensional space relative to the current location of one's own body (or parts thereof). E.g., in order to pick a peach it is not sufficient to have allocentric information about the peach's location relative to the rest of the scene, you must also have egocentric information about its position relative to the current location of your hand. It is in this sense, to a first approximation, that both perception and action make use of a common, egocentric way of representing objects in space.

As an aside, I should mention that it does not follow from the fact that it is possible to specify an object's perceived position in egocentric space using the subject-centered axes *left/right*, *above/below*, and *in front of/behind* that a spatial coding system based on these axes is actually used to plan actions directed in relation to the object. Indeed, behavioral and neurophysiological studies rather suggest that visuomotor activity is typically subserved by a variety of coordinated, effector-specific spatial coding systems, some of which represent object locations not in extrinsic, directional terms, but rather in intrinsic, kinematic terms (Scott 2008, Kalaska 2009).

---

same time. The idea is rather that, when she perceives an object's position in visual egocentric space, it may be any part of her body of which she is proprioceptively aware in relation to which the object's position is perceived. See Briscoe 2009: 425-426.

[47] Although I think that a verdict here should be based, in part, on a theory's ability to explain the phenomenon of *visual direction constancy*. E.g., Peacocke must explain why the tree's direction appears stable when I maintain fixation on the tree, but rotate my torso to the right, if the tree's unchanging position relative to my *eyes* is not also represented in visual experience. See Wu 2014 for a recent argument that the experience of visual direction constancy requires an encoding of perceived object locations in one or more non-retinocentric, egocentric frames of reference.

Tim Crane (2009) points out that it does not follow from the fact that it is possible to describe the content of a perceptual experience using a proposition that the inner representational vehicle of that experience is itself propositional or sentence-like. I am making a similar point here about egocentric spatial representation. We need to distinguish between a specification of the objective content of an egocentric spatial representation *R*, which may be characterized using any geometrically adequate system of spatial representation, and a specification of the system of spatial representation actually used to construct *R*.

Let's turn to Claim 3. It deals, importantly, with the relationship between what we see and what we do. One way of interpreting Claim 3 would be as insisting that the egocentric spatial content of a visual experience is not constitutively independent of its *motoric significance* for the perceiving subject. By the motoric significance of a visual experience, I intend its functional role in planning spatially directed bodily actions in light of the agent's current motivations, beliefs, objectives, etc. There is a clear connection between this way reading Claim 3 and Jesse Prinz's recent proposal that conscious perceptual experience functions as a *menu for action*: "Consciousness makes information available for decisions about what to do, and it exists for that purpose" (Prinz 2012: 203).

Two remarks are important. First, Claim 3 connects having spatially significant perceptual information with having "*dispositions* to do various things." Why? One thought is that Evans intended to establish certain general, non-intentionally characterizable necessary conditions for perceptual spatial representation (see, e.g., Noë 2004, chap. 3). I think that this behaviorist interpretation of Evans's project in chapter 6 of *The Varieties of Reference* is mistaken. Evans, in several places, is explicit that dispositions to spatial behavior need not be directly induced by perceptual inputs, but may be, as he says, "conditional also on other beliefs and desires" (1982: 155, fn. 23). A subject, capable of reasoning, will be disposed to engage in an action directed in relation to certain a region of egocentric space only "when [her] thoughts make it appropriate" (1982: 161, fn. 33) or when it seems a "good thing to do" (1982: 161). She will not normally form a disposition to advance in the perceived direction of angry rattlesnake or to reach for a red-hot poker, for example, unless her circumstances provide her with compelling *reasons* to do so. In general, Evans' dispositions to spatial behavior appear to be propensities to respond to the spatial structure of the environment in ways that are intelligible in light of the subject's cognitive and motivational states. They are not dispositions to behavior in the sense of the classical behaviorist. What matters, for Evans, is knowing how to act in response to the perceptual inputs in ways that make rational sense.

Second, it is important not to collapse the distinction between the egocentric spatial content of a visual experience and its motoric significance. Evans's view is not that the egocentric spatial content of a perceptual experience just *is* its "behavioral spatial purport," to use Rick Grush's (2007) term. It is one thing to specify the egocentric spatial content of a visual experience, say, that an object is located in a certain direction and at a certain distance in depth from the subject. It is quite another to provide an account of what makes it the case that, for any given egocentric spatial content, the subject is having a visual experience with *that* content and not some other.[48] Evans's proposal is that such an adequately individuating account of the egocentric spatial contents of perceptual experience must advert to the subject's abilities to engage in bodily actions targeted on or otherwise directed in relation to environing objects and surfaces. Motoric significance is that in virtue of which visual experiences have objective, egocentric spatial content. This proposal, however, should not be taken to imply that the egocentric spatial content of a visual experience is identical to its motoric significance. The latter is not *what* we see when we have a visual experience with a certain spatial content. Rather, it is part of *having* a visual experience with that content and not some other. (Compare the proposal that a statement *S* must play a certain inferential role in order to have the semantic property of being true just in case *Edinburgh is east of Glasgow, but west of Oxford*. This proposal clearly does not imply that *S*'s truth-condition is in any sense identical to its role in inference.)

This, I take it, is the point Evans is making with Claim 4:

> …when I speak of information "specifying a position in egocentric space," I am talking not of information about a special kind of space, but of a special kind of information about space—information whose content is specifiable in an egocentric spatial vocabulary. It is perfectly consistent with the *sense* I have assigned to this vocabulary that its terms should *refer* to points in a public three-dimensional space (1982: 157).

The claim that the egocentric spatial content of a visual experience is determined by its motoric significance for the perceiving subject thus is a claim not about *what* the visual experience represents (it is not a claim at the level of reference), but a claim about the conditions under which a visual experience will have such content. When a subject perceives an object's egocentric location, the information about space in her possession is "special" because it is poised to guide her actions in relation to the object. It is consistent with this suggestion, however, that egocentric spatial properties represented in her perception are objective (if subject-relative) properties of things in public, three-dimensional space.

[48] This formulation is the perceptual analogue of Peacocke's "Discrimination Principle" for the conceptual contents of thought.

A number of philosophers have failed adequately to enforce the distinction between egocentric spatial content and what I am calling motoric significance. For example, Adrian Cussins, in a discussion of the spatial content of auditory perception, writes:

> Evans's idea is that the spatial content of the auditory perception has to be specified in terms of a set of conceptually unmediated abilities… to move in the egocentric space around the organism. This is because *the content consists* in the experiential availability to the subject of a dispositional ability to move. The experiential content of perception is *specified in terms of certain fundamental skills which the organism possesses* (1990: 397, emphasis added).

One natural objection to saying that the egocentric spatial content of a perceptual experience consists in having certain abilities to move one's body is that there are indefinitely many different ways in which a perceiver may elect to respond to the perceived spatial structure of the environment. As Mohan Matthen writes, "there is no such thing as *the* proper response, or even a range of functionally appropriate responses, to what perception tells us" (1988: 20). The egocentric spatial content of a visual experience, for Evans, however, does not consist "in the experiential availability to the subject of a dispositional ability to move." Motoric significance is that *in virtue* of which a visual experience has egocentric spatial content. It is not identical to (the same thing as) its egocentric spatial content.

John Campbell (2005) also seems to elide the egocentric spatial content of a visual experience with its motor significance in a discussion of Evans. According to Campbell, Evans's egocentric spatial representations identify the locations of objects "merely as affordances," as possible ways of moving and acting (2005: 200). Campbell rightly balks at this proposal: "The trouble with this gloss on the content of egocentric identifications of location is that we would ordinarily take spatial location to be the *categorical basis* of these affordances. That is, we think that it is the relative locations of the thing and the agent that explain why it is possible for the agent to act on the thing. We do not suppose that egocentric location is actually constituted by the possibility of the agent acting on the thing" (2005: 201).

Evans, however, on the interpretation offered here, does not collapse the distinction between what we see and what we do in this way. The point to emphasize is that egocentric spatial information is not information about a special kind of space, a space of Gibsonian affordances, but rather information of a special kind about space, information poised to guide actions that are sensitive to the spatial properties of the objects around us. The claim that "having spatially significant perceptual information consists at least partly in being disposed to do various things" (Evans 1982: 155) does not imply that such information is constituted by (or about) what one is disposed to do. It is *having* spatially significant perceptual information that, for Evans, partly consists in being disposed to do various things, not the information itself. As Jesse Prinz has forcefully argued, perceptual experience from a functional standpoint can be for action without being constituted by action (Prinz 2012: 211).

Last, if a "sensorimotor chauvinist," in Andy Clark's (2009) sense, is someone who maintains that any difference in the motor activities to which a perception may give rise constitutes a difference in the perception's content, then it is clear that Evans is not a sensorimotor chauvinist. He respects the distinction between the content of a representation and the use to which that representation is put by its motoric consumers. One implication is that different perceivers (even those belonging to different species) may, in principle, have visual experiences with the same egocentric spatial contents despite having very different motor skills. Thus, I and an orangutan may both see that a tree branch is, e.g., at certain orientation in depth even though the motoric significance of the orangutan's visual experience is no doubt very different than that of my own visual experience. It is a significant merit of the present interpretation of Evans's theory, I take it, that it enables Evans to sidestep the charge of sensorimotor chauvinism sometimes leveled against action-oriented approaches to the spatial contents of perceptual experience.

3. Looking at Action-Oriented Represenation from a Biosemantic Perspective

Ruth Millikan's biosemantic theory of mental representation (Millikan 1989, 1995, 2004), I would suggest, provides us with resources for developing Evans' approach with a bit more detail. According to the biosemantic theory, an item R_0 will function as a representation of some structured aspect of the distal environment E_0 only if two conditions obtain:

> *Guidance*: R_0's function or purpose is to guide a consumer in the performance of some type of task T (or a range of different types of tasks), where the consumer's successful performance of T depends on the fact that E_0 obtains.
>
> *Systematicity*: The way the consumer is guided by R_0 systematically depends on R_0's structure or composition, such that had some variant of R_0 (R_1,... R_n) been produced instead of R_0, then the consumer's way of performing T would have proved successful only if instead of E_0 there had been some corresponding variant (E_1,... E_n).

The idea that perceptual awareness of viewer-relative spatial layout constitutively involves a kind of bodily readiness for action—Claim 3 above—can be elaborated using versions of these two conditions. Assume that R_0 was produced by a perceptual input system of some kind. Then, R_0 will represent the instantiation of a determinate egocentric spatial property G_0, e.g., a certain viewpoint-relative distance or direction, by an object *O* only if:

> *Motor Guidance*: R_0's function or purpose is to guide motor consumers in the performance of some type of action *A* (or a range of different types of actions), where successful performance of *A* depends on the instantiation of G_0 by *O*.
>
> *Motor Systematicity*: The way motor consumers are guided by R_0 systematically depends on R_0's structure or composition, such that had some variant of R_0 (R_1,... R_n) been produced instead of R_0, then their way of performing *A* would have proved successful only if instead of instantiating G_0 some corresponding determinate of the same determinable property (G_1,... G_n) had been instantiated by *O*.

We can consider these two requirements in relation to a toy model of how connectionist neural networks handle the problem of sensorimotor coordination devised by Paul Churchland (1986, 2012). The model involves a virtual, robotic crab with a two-jointed moveable arm and eyes that can rotate 90° from side to side (**Figure 1**). The crab represents the location (*x*, *y*) of an object in front of it using a sensory activation vector (α_i, β_i), where *x* and *y* specify the object's placement in a two-dimensional coordinate system centered on the hinge of the crab's shoulder and where α_i and β_i are the input activation levels corresponding, respectively, to the rotation angles α and β of its left and right eyes.

The crab's task is to move its arm so as to position the tip of its pincer on the (*x*, *y*) coordinates of a reachable object *O*. That arm position will require the crab's shoulder and elbow to assume a pair of angles (θ, ϕ). For example, if *O* is located at the point encoded by eye-angles (62°, 98°), then the crab's shoulder and elbow joints must assume the angle pair (60°, 47°). In order to perform this task, the crab's neural network is trained to transform the eye-angle input vector (α_i, β_i) into a motor output vector (θ_o, ϕ_o), where θ_o is output activation level corresponding to the shoulder angle θ and ϕ_o is the output activation level corresponding to the elbow angle ϕ.

It is easy to see that Churchland's crab meets the *Motor Guidance* requirement: a given eye-rotation-angle activation pair in the crab's input layer (α_i, β_i) represents the distal location (*x*, *y*) of an object *O* because the vector's *function* is to guide the way the crab reaches toward *O*, and the crab will only reach toward *O* successfully only if *O* is situated at (*x*, *y*). The crab also meets the *Motor Systematicity* requirement: For any reachable location (*x*, *y*), there is a corresponding vector (α_i, β_i)) in the crab's sensory input layer. To each such input vector, in turn, there is a corresponding vector (θ_o, ϕ_o) in the crab's motor output layer that will guide the tip of the crab's pincer to (*x*, *y*). In other words, variations in *O*'s distance and direction systematically give rise to variations in sensory input which, in turn, systematically give rise to variations in object-directed motor output. There is a one-to-one mapping from points in objective space to points in visual space to points in motor space.

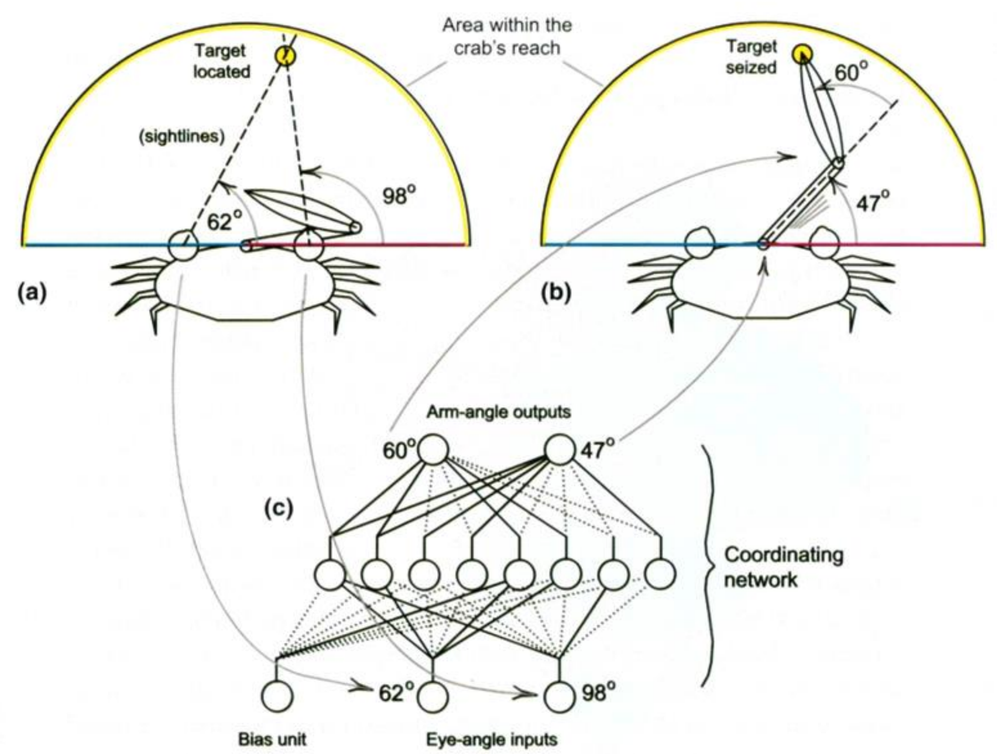


Figure 1. A robotic crab with an extendable arm and rotatable eyes
From Paul Churchland. 2012. *Plato's Camera: How the Physical Brain Captures a Landscape of Abstract Universals*. Plate 3 © Massachusetts Institute of Technology, by permission of The MIT Press.

I will close with a few brief remarks about this example. First, the robotic crab example is deliberately simple. Besides having more sophisticated perceptual systems, real-world sensorimotor agents typically have a wide array of controllable body parts, access to proprioceptive/kinaesthetic information concerning their movements, and the ability to select goals and types of actions to perform on the basis of their beliefs and current needs.

When we scale up, in particular, when the way consuming motor systems respond to inputs from perceptual producers in guiding actions depends, in part, on belief- and desire-sensitive practical reasoning, there will be no simple relationship between sensory inputs and motor outputs. Hence, as Matthen says, there will be no such thing as *the* functionally proper response to what is perceived. (Indeed, much of the time, the correct response to the representation will be to do nothing at all, to ignore the bit of reality that the representation reveals.)

This is not a problem for the approach sketched here. For purposes of content-individuation, what is important from the standpoint of the biosemantic interpretation of Evans' theory is not uniformity in the *effects* a perceptual representation has on its consumers, but rather uniformity in the worldly condition under which consumer responses, however diverse these may be, will prove successful. As Millikan writes, "if the position of the chair in the room does not correspond, so, to my visual representation of its position, that will hinder me equally in my attempts to avoid the chair when passing through the room, to move the chair, to sit in it, to remove the cat from it,… etc." (Millikan 1989: 289).

Perception and action, then, although less closely linked, are not decoupled in sophisticated sensorimotor agency. For any determinate, spatial property *G*, there will be indefinitely many different ways in which the agent might respond to the experience of *G*'s instantiation by an object in her field of view. What matters to Evans' approach, as reconstructed in this section, however, is not *which* of various possible actions the agent actually selects for performance. Rather, what matters is that the agent's visual experience is poised to guide a range of actions directed in relation to the object and that the *way* any given action in the range is performed depends for its success on *G*'s being instantiated. When an experience is poised in this way, it equips the agent with the practical know-how needed to interact with the object in *G*-sensitive ways.

The second point is that such know-how need not be implemented in order to have the experience in question. Having the experience is dependent on having the *capacity* to perform *G*-sensitive actions rather than on its overt actualization (see Schellenberg 2007 for discussion of this idea). What matters to seeing an object's direction, for example, is not acting on one's capacity to walk (or run or crawl) in its direction, but knowing how one *would* have to move one's body in order to do so. Evans' theory does not have the implausible implication that an agent who is paralyzed or unable to move her body—say because she has been buried up to her neck in sand—is thereby unable perceive the spatial attributes of the objects that surround her. So long as she has the right sort of practical know-how, she qualifies as a fully-fledged space perceiver.

Third, the distinction between an experience's objective spatial content and its motoric significance for the perceiving subject (Claim 4) can be clearly drawn within the biosemantic framework: the spatial content of a given perceptual representation—in the crab example, this representation will be a given sensory input vector (α_i, β_i)—is not identified with the functions or activities performed by its "downstream" motoric consumers, with what its motoric consumers *do*. Rather, it is identified with the way the world needs to be if the representation's motoric consumers are to perform their functions successfully—here the way object *O* must be situated in front of the crab if the $(\alpha_i, \beta_i) \rightarrow (\theta_o, \phi_o)$ transformation is to result in successful reaching.

It is thus possible to maintain that perceptual representations acquire spatial significance in virtue of their functional role in guiding actions without collapsing the distinction between the way the distal world is represented as being in perception and the subject's motoric responses to the world. Objects are not represented merely as things that can be acted upon thus-and-so, but as having the intrinsic and relational spatial properties that afford possibilities for action.

Finally, human perceptual systems produce representations that are not limited in their function to action-guidance, but that also play a role in high-level object recognition, imaginative problem-solving, and intersubjective communication. As Noë argues, it would thus be "dogmatic to suppose that the *only* aim of vision is action" (2010: 248). The *Motor Guidance* requirement, however, does not entail commitment to this dogmatic supposition. To insist that the spatial content of a perceptual state derives from the state's functional role in action-guidance is not to deny that the state may serve a wide variety of other purposes as well. The *Motor Guidance* requirement in no way precludes a role for space-representing perception in forms of problem-solving that are related only in a very indirect ways to negotiating and interacting with the three-dimensional environment. Perceptual experience, from a functional standpoint, can be *for* action without being exclusively for action.

4. Conclusion

There is obviously much more that needs to be said about and in defense of Evans' approach to the relationship between action and perception. I have not said anything, in particular, about empirical objections to the approach premised on the dual systems model of visual processing (Milner & Goodale 1995/2006; Clark 2001, 2007; Goodale & Milner 2004).[49] I hope however to have shown in this brief treatment that there is a coherent alternative to actionism, one that looks to embodied *interaction* with the distal environment rather than to knowledge of the sensory consequences of movement to explain the spatial contentfulness of perceptual experience.[50]

[49] But see Briscoe 2008b, 2009 and Briscoe & Schwenkler *forthcoming*.

[50] For helpful discussion of an earlier version of this paper, I am grateful to James Genone and Wayne Wu.

References

Arterberry, M. 2008. Infants' sensitivity to the depth cue of height-in-the-picture-plane. *Infancy*,13: 544-555.

Berkeley, G. 1709/2008. *An Essay towards a New Theory of Vision*. In D. Clarke, ed. *Berkeley: Philosophical Writings*. Cambridge: Cambridge University Press.

Bermúdez, J. 1998. *The Paradox of Self-Consciousness*. Cambridge, MA: MIT Press.

Bermúdez, J. 2005. The phenomenology of bodily awareness. In D. W. Smith and A. Thomasson, eds. *Phenomenology and Philosophy of Mind*. Oxford: Oxford University Press.

Bermúdez, J. 2007. What is at stake in the debate on nonconceptual content? *Philosophical Perspectives*, 21: 55-72.

Block, N. 2005. Review of *Action in Perception*. *Journal of Philosophy* 102: 259-272.

Block, N. 2003. Tactile sensation via spatial perception. *Trends in Cognitive Sciences* 7: 285-286.

Block, N. 2012. Discussion of J. Kevin O'Regan's *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness*. *Review of Philosophy and Psychology* 3: 89-108.

Braddick, O. & Atkinson, J. 1983. Some recent findings on the development of human binocularity: A review. *Behavioural Brain Research* 10: 141-150.

Briscoe, R. 2008a. Vision, action, and make-perceive. *Mind and Language* 23: 457-497.

Briscoe, R. 2008b. Another look at the two visual systems hypothesis. *Journal of Consciousness Studies* 15: 35-62.

Briscoe, R. 2009. Egocentric spatial representation in action and perception. *Philosophy and Phenomenological Research* 79: 423-460.

Briscoe, R. forthcoming. Bodily action and distal attribution in sensory substitution. To appear in *Sensory Substitution and Augmentation*, edited by F. Macpherson, *Proceedings of the British Academy*, Oxford: Oxford University Press.

Briscoe, R. and Grush, R. in preparation. Action-based theories of perception. To appear in the *Stanford Encyclopedia of Philosophy*.

Briscoe, R. and Schwenkler, J. forthcoming. Conscious vision in action. To appear in *Cognitive Science.*

Burge, T. 2005. Disjunctivism and perceptual psychology. *Philosophical Topics* 33: 1–78.

Burge, T. 2010. *Origins of Objectivity.* Oxford: Oxford University Press.

Burge, T. 2005. Information-processing, phenomenal consciousness, and Molyneux's problem. In José Bermúdez, ed., *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*. Oxford: Oxford University Press.

Campos, J., Bertenthal, B., & Kermoian, R. 1992. Early experience and emotional development: The emergence of wariness of heights. *Psychological Science* 3: 61–64.

Campos, J., Anderson, D., Barbu-Roth, M., Hubbard, E., Hertenstein, M., & Witherington, D. 2000. Travel broadens the mind. *Infancy* 1: 149-219.

Cassam, Q. 1997. *Self and World*. Oxford: Oxford University Press.

Churchland, P. 2012. *Plato's Camera.* Cambridge, MA: MIT Press.

Clark, A. 2001. Visual experience and motor action: Are the bonds too tight? *The Philosophical Review* 110: 495-519.

Clark, A. 2007. What reaching teaches: Consciousness, control, and the inner zombie. *The British Journal for the Philosophy of Science* 58: 563-594.

Clark, A. 2009. *Supersizing the Mind.* Oxford: Oxford University Press.

Clark, A. 2012. Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind* 121: 753-771.

Crane, T. 2009. Is perception a propositional attitude? *Philosophical Quarterly* 59: 452-469.

Cussins, A. 1990. The connectionist construction of concepts. In M. Boden, ed., *The Philosophy of Artificial Intelligence.* Oxford: Oxford University Press.

Cutting, J. & Vishton, P. 1995. Perceiving layout and knowing distances. In W. Epstein and S. Rogers, eds. *Perception of Space and Motion.* San Diego, CA: Academic Press, 69–117.

Evans, G. 1982. *The Varieties of Reference*. Edited by J. McDowell. Oxford: Oxford University Press.

Evans, G. 1985. Molyneux's question. In A. Phillips, ed. *The Collected Papers of Gareth Evans* London: Oxford University Press.

Ganz, L. 1975. Orientation in visual space by neonates and its modification by visual deprivation. In A. Riesen, ed. *The Developmental Neuropsychology of Sensory Deprivation.* New York: Academic Press, 169-210.

Goodale, M. 2011. Transforming vision into action. *Vision Research* 51: 1567-1587.

Granrud, C. & Yonas, A. 1984. Infants' perception of pictorially specified interposition. *Journal of Experimental Child Psychology* 37: 500-511.

Granrud, C., Haake, R., & Yonas, A. 1985. Infants' sensitivity to familiar size: The effect of memory on spatial perception. *Perception and Psychophysics* 37: 459–466.

Gunther, Y. 2003. *Essays on nonconceptual content*. The MIT Press.

Grush, R. 2007. Skill theory v2.0: Dispositions, emulation, and spatial perception. *Synthese* 159: 389-416.

Held, R., & Hein, A. 1963. Movement-produced stimulation in the development of visually guided behavior. *Journal of comparative and physiological psychology*, *56*: 872-876.

Kalaska, J. 2009. From intention to action: Motor cortex and the control of reaching movements. In D. Sternad, ed., *Progress in Motor Control*. Berlin: Springer.

Kellman, P. & Arterberry, M. 2000. *The Cradle Of Knowledge: Development of Perception in Infancy*. Cambridge, MA: MIT Press.

Kellman, P. & Arterberry, M. 2006. Perceptual development. In W. Damon, D. Kuhn, & R. Siegler, eds. *The Handbook of Child Psychology: Cognition, Perception, and Language*. Hoboken, NJ: Wiley, 109–160.

Matthen, M. 1988. Biological functions and perceptual content. *The Journal of Philosophy 85*: 5-27.

Matthen, M. 2006. Review of *Action in Perception*. *Mind* 115: 1160-1166.

McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.

Millikan, R. 1989. Biosemantics. *Journal of Philosophy* 86: 281-297.

Millikan, R. 2004. *Varieties of Meaning*. Cambridge, MA: MIT Press.

Millikan, R. 2006. Postulating perceptual representations in a way that actually supports Gibson's central insights. Lecture at the Boston Colloquium for Philosophy of Science, January 30, 2006.

Milner, A.D. and Goodale, M. 1995/2006. *The Visual Brain in Action*, 2nd edition. Oxford: Oxford University Press.

Náñez, J., & Yonas, A. 1994. Effects of luminance and texture motion on infant defensive reactions to optical collision. *Infant Behavior and Development* 17: 165-174.

Noë, A. 2004. *Action in Perception.* Cambridge, MA: MIT Press.

Noë, A. 2010. Vision without representation. In N. Gangopadhyay, M. Madary, & F. Spicer, eds. *Perception, Action, and Consciousness*. Oxford: Oxford University Press, 245-256.

O'Regan, J. K. 2011. *Why Red Doesn't Sound Like a Bell: Explaining the Feel of Consciousness* Oxford: Oxford University Press.

O'Regan, J. K. and Noë, A. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences,* 24: 939–1031.

Peacocke, C. 1992. *A Study of Concepts*. Cambridge, MA: MIT Press.

Prinz, J. 2006. Putting the brakes on enactive perception. *Psyche*, 12: 1-19.

Prinz, J. 2012. *The Conscious Brain*. Oxford: Oxford University Press.

Schellenberg, S. 2007. Action and self-location in perception. *Mind* 116: 603-631.

Scott, S. 2008. Inconvenient truths about neural processing in primary motor cortex. *Journal of Physiology* 586: 1217-1224.

Wu, Wayne 2014. Against division: Consciousness, information and the visual streams. To appear in *Mind and Language.*

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0010

Naturalizing Husserlian Phenomenology along a Leibnizian Pathway[51]

Jean-Luc Petit
Université de Strasbourg
jean-luc.petit[]college-de-france.fr
Received December 2013; accepted August 2014; published Autumn 2014.

Abstract

A contribution to the history of a formerly hotly discussed, but short-lived scientific project: neurophenomenology[52], the proposal of weaving together Husserlian phenomenology of consciousness and the neuroscience of brain functioning, this article traces back the opening and closing of an apparent window of opportunity, both in phenomenology and in neuroscience, for the eventually unfulfilled realization of that project.

Keywords: neurophenomenology; phenomenology; coherence; *autopoiesis;* constitution.

Despite a growing literature on naturalizing phenomenology, the subject matter remains controversial. That might be because phenomenology is not intended like natural science to discover a class of facts—such as facts relative to conscious states of human mind. The purpose of phenomenology is rather to reveal the world's mode of appearance to a perceiving subject. In this paper I will put to the test the chances of a recent trend in neuroscience to stand as a satisfactory candidate program for naturalizing a brand of phenomenology at first sight irreducible to natural science: Husserl's transcendantal constitution theory.

1. Mixed perspectives for naturalizing phenomenology.
2. *„Konstitution durch Einstimmigkeit der Erfahrung"*.
3. Brain orchestra *vs* single-cell machine-gun .
4. Coherence by way of *autopoiesis*.
5. In-brain and between-brain coherence.
6. A challenge to pheno-physical reductionism?

[51] Acknowledgement to Dr Christopher Macann for the translation.

[52] On the subject of connection between neurophenomenology and enactivism, see note at the end of the article. [eds.]

1. Mixed perspectives for naturalizing phenomenology

At the grand congress *Actualité cognitive de la Phénoménologie: Les défis de la naturalisation* in Bordeaux (19-21 October 1995), the official launch of a cognitive science project for naturalizing Husserlian phenomenology (bridging the gap separating it from natural science), to clarify matters in the discussion I had the happy inspiration of drawing on a transparent a vertical wavy line, on either side of which I put designations of brain processes, and facing each the phenomenological structures of experience I pretended to be "the same thing under another name". Incidentally, that was the very first expression in a philosophical environment of the hunch—it would assume later the proportions of a scientific hypothesis—that mirror neurons recently identified by Rizzolatti and his team might be the brain substrates of *Einfühlung* as conceived by Lipps and Husserl. The wavy line was soon retaken on their behalf as code name of the naturalization problem at large by such luminaries as Francisco Varela, Bernard Pachoud, Jean-Michel Roy and Jean Petitot, so that my suggestion was pushed aside and reduced to the triviality of "simply to draw the parallel and leave it unexamined", while they themselves were in the serious business of building the *neurophenomenology* that the same Varela had heralded (Petitot et al. 1999: 66-67). If you do not mind that I resume a dialogue death has interrupted, I would like to recap my point: 1°) systematically correlating constitutive operations in Husserlian phenomenology with functional mechanisms in neuroscience is both meaningful and feasible; 2°) far from being a fortuitous analogy, such correlation relies on a fundamental identity of nature: basically it is indeed the same thing; 3°) however, lacking a common language to determine that identity, between the subjectivism of phenomenological description and the objectivism of explanation in terms of brain mechanisms the dualism—perhaps only a semantic dualism—is probably insurmountable. But it would be mean of me to move the target after the shot was fired. In this paper I would prefer to examine the chances of another possible way of bridging the gap between transcendantal constitution theory and brain functional systems. That way, without having been followed by Husserl, was at least left open by him as a possibility that his choice of expressions convincingly suggests. And to prove that I am not mean towards Varela I am willing to concede that he spawned the trail in question, if only as a limited application of his ambitious program of neurophenomenology. This time I anchor my proposal in a common language on both sides of the empirical-transcendental divide. I will capitalize on the fact that instead of two incommensurable discourses, use is made both in phenomenology and neuroscience of a multi-scale concept; I mean a concept applicable to different levels of organization of the living being: *coherence*. One might speak of coherence at the personal level of the perception and behavior of a subject; but equally at the sub-personal level of assemblies of neurons in the brain. Of course, the question remains how it is possible to move from coherence from the point of view of a subject to coherence internal to the brain metabolism: a coherence without a subject, which, however, is not an absolute property of brain anatomical structures, neither might possibly enjoy full autonomy with respect to the material substrate of the brain, but needs to be eventually relativized to functional loops and other rhythms of the organism, as somatic precondition to a coherent subjective experience.

Well-known is the tribute paid to Descartes dualism by transcendantal phenomenology; less so is the growing relevance to Husserl's later oeuvre of Leibniz monadology, a hierarchically organized plurality of monads whose unfolding and refolding are harmoniously synchronized. If Cartesian dualism hinders the circulation of sense between the levels of organization of the living, the ubiquity of sense is the rule in monadological context. In a manuscript from 1935 Husserl warned against localizing mental life in the brain. He relied on an analogy between the organism as a society of cells and society at large. If the intersubjective life of society is not to be tied to the bodies of participants, then the subjective life of a man should not be attached to the society of his bodily cells. However, in an earlier text from 1929 Husserl expressed less negative views while toying with Leibniz' idea of monad as a way of founding subjective life on the system of monadic cells of the organism. Accordingly, Leibniz' favourite metaphor of non-causally interacting synchronized clocks might pave the way to naturalization of phenomenology.

2. *„Konstitution durch Einstimmigkeit der Erfahrung"*

Rejecting every prior ontological assumption, Transcendantal Constitution Theory is uniquely reliant on the intrinsic resources of the perceiving and acting subject of endowing with a sense of being privileged episodes of his experience. *Einstimmigkeit der Erfahrung:* the tuning of experience with itself proves to be a purely immanent criterion of sense-giving, regardless of the level of organization of experience. A perceiver is prone to intentionally aim at an object in the world as soon as he notices a tuning where there was only discordance: whether between visual field sketches (*Abschattungen*) and courses of kinaesthesia of the moving organs, between the various sensory modalities, between the body as physical object and the body as lived from within, between the world as viewed by oneself and by others, a tuning he may eventually project at the limit of all possible experience—as in physics—under the Idea of Nature. Typical of Husserl's transcendental phenomenology is the constant back-reference of the constitutive operations of giving sense of being to the transcendental subject as their bearer. Having sense for... is a relational, not an absolute property. For any entity in experience to be posited as a reality in the world, it needs a special constitutive act of the subject. Having said that, to constitute does not mean to create. The fact that I implement the constitutive operations makes me responsible for the meaningfulness of my experience, but does not turn me into a small creative god. Holding at bay the threat of a solipsistic illusory arbitrariness, the course of our experience harbors a necessity that excludes any caprice even if that necessity is not that of logic and if it manifests itself only in a gradual way. This necessity consists in that even if the discordance is always possible—even if it occurs from time to time—its impact is always limited and does not preclude a more general concordance to recover. Of course, one can always try to assign the privilege of concordance over discordance to the subject, whether the cognitive subject grounding inferences to the future on the basis on prior experience, or the perceiving subject anticipating the constancy of objects in a stable world, or the moral subject who wants to give meaning to his existence and achieves this end in action. Only that we quickly realize that we are caught in a turnstile: on the one hand, each form of concordance in experience should be credited to the constituting subject; on the other hand, the constitutive act of the subject escapes arbitrariness only in relying on the concordance of experience. Is there a way out of that impasse? At first sight it would be like reconciling the materialist monism of neuronal explanation with the dualism of psycho-physical causation. Might concordance possibly emerge from brain metabolism as if summoned to existence by the constitutive power of the transcendental subject, while dispensing with the transcendental stance?

3. Brain orchestra *vs* single-cell machine-gun

Individual cells in the brain are spontaneous oscillators. Specifically the neurons in early visual areas emit a sudden burst of electrophysiological discharge: a potential of action, provided that the preferred stimulus entered their receptive field, the portion of visual field they care for exclusively. A series of potentials of action alternating peaks and troughs gives a wave of electrical activity selectively associated with the stimulus in question and that electrophysiologists see it as a code of the recognition by the perceiving organism of the corresponding aspect of the environment. The propagation of action potentials through synaptic connections from neuron to neuron along the hierarchically organized pathway dedicated to the cognitive treatment of external information obeys the same pattern. It all boils down to a business of anatomical localization and electrical power: a pure matter of energy accumulation somewhere in the brain and energy expenditure for carrying elsewhere the relevant information following a strictly predetermined path of cortical wiring in the white matter of brain tissue. There is no possibility of an action at a distance whether in the brain or between brains, an action, that is to say, an influence without a definite material carrier such as the sensory and motor pathways, an influence from the center to the periphery, from the global to the local, from top to bottom. Such a view of things makes a mystery of performances as unexceptionable as orienting one's attention towards an object of interest in the surroundings, conceiving a motor intention and mobilizing the necessary resources to carry out the purpose, or even making oneself understood by the receiver of one's message.

In 1958 Hubel and Wiesel discovered the preferred stimulus of a neuron in the visual cortex of cats and monkeys: they were trying to stimulate a cortical cell by their mentor Stephen Kuffler's method of using slides to project spots onto a screen in front of the animal. Hubel explains:

> Then gradually we began to elicit some vague and inconsistent responses by stimulating somewhere in the midperiphery of the retina. We were inserting the glass slide with its black spot into the slot of the ophthalmoscope when *suddenly over the audiomonitor the cell went off like a machine gun*. After some fussing and fiddling we found out what was happening. The response had nothing to do with the black dot. As the glass slide was inserted its edge was casting a faint but sharp shadow, a straight dark line on a light background. That was what the cell wanted, and it wanted it, moreover, in just one narrow range of orientations. (Hubel & Wiesel 2004: 661)

A significant change of paradigm took place in neuroscience when the search for single-cell coding of an elementary trait of the environment by a selective burst of electric activity gave way to an investigation of the dynamics of long-distance interactions between brain regions subtending a complex cognitive function. For assemblies of neurons to communicate regardless of their location in the brain, it suffices that their oscillations be transitorily in-phase (phase-locking), a condition requiring no supplementary amount of energy on top of the component neurons activities. Let's consider such in-brain communication by temporal coherence rather than energy consumption or spatial distribution as the nearest possible analogue of Husserl's tuning-based constitution.

4. Coherence by way of *autopoiesis*

On the paradigm shift in neuroscience Varela's ideas have had a profound impact—even if not always acknowledged nowadays. To recap an early article: "Patterns of Life: intertwining Identity and Cognition", dating back to 1997, central to the autonomy of the living being is *autopoiesis*: a process by which the living being ensures its own production, uniquely characterized by the emergence of a *coherence* of some kind. The interesting fact about coherence is that while being an effective source of interaction, it is minimally dependent on energy flow, and needs neither controller nor a fixed localization. An autopoietic process, coherence is—so to say—self-produced: it will maintain its organization as long as its basic process resists perturbations and will dissolve when confronted with perturbations that go beyond its viability margins. In the background of such ideas, it is no wonder if Varela's last work applied to emergence of neural assemblies by coherence or synchrony of oscillations, as it is perhaps the simplest, and less objectionable manifestation of autopoiesis in the living being. Let's be more specific about coherence in neuronal networks as an autopoietic mechanism. An individual neuron's action potential is a signal to be understood both ways, as energy expenditure and as information. In visual neurons the information is the noticing of detection of the presence of a preferred stimulus in the cell's receptive field. Might information be decoupled from energy and location? It seems so. A transitorily occurring coherence of the oscillatory activities of distant neuronal groups in the mess of different rhythms in the brain might also convey information about the fact that a new neuronal assembly has just been formed by the grouping of the component neuronal groups through phase coincidence of their respective oscillations. Such information might be relevant to the system in case that neuronal assembly subtended some cognitive function. Yet, the information requires no additional power supply besides that which is consumed by the action potentials of individual neurons. More precisely, energy in the brain is measured by the change in phase amplitude of oscillations: a deflection of electric curve in relation to base-line, a blend of spatial cum intensity dimension of neural activity. In contradistinction, phase synchrony uniquely measures the temporal relationships between neuronal group activities, independently of their amplitude. Two signals cohere with each other if their alternating spikes and troughs succeed synchronously. The constancy of phase similitude (or phase difference) of the two signals during a short lapse of time suggests the existence of a mechanism capable of locking together (phase-locking) at a distance the paces of deployment of both events. The resulting transitorily stable activity pattern over distant brain areas would open up a window of communication between these areas that owes little to the discharge rate of the underlying neurons. Such disassociation between the temporal dimension of brain activity and its dimensions of energy and location opens up a golden avenue for speculations about the possibility of capturing—at a higher level of the description of neural dynamics if not at the fine-grained level of individual neurons—the evidence of a purely informational and non-material nor causal interaction between the systems (or subsystems) of the living being. Which not only brings us back to autopoiesis, a mode by which the organism persists in being alive despite the changing energy states in the physical environment, but beyond naturalization of constitution theory by autopoiesis, such move would tend up to legitimize rather uncritically what amounts to an interactionist mind-brain dualism à la Descartes, except that any reference to the thinking subject is discarded in favor of a non-matter-energy characterization in purely temporal terms of the neuronal basis of cognitive behavior.

In a famous research by Varela and his LENA team at La Salpêtrière (Rodriguez et al, 1999) participants were presented with heavily contrasted photos indistinguishable from random shapes when presented upside down but easily recognizable as human faces when presented upright (Mooney figures). Focusing on whole brain γ band electric activities (40 Hz) induced by the stimulus and recorded through electrodes placed on the scalp, they first computed maps of phase amplitude variations of oscillations showing the two expected peaks of spectral power related to presentation of stimulus and to motor response. But they also computed a dynamic mapping of the spatio-temporal distribution of γ activities on the scalp from stimulus presentation to motor response, showing that while γ activity stayed relatively homogeneous in power emission in all conditions, phase temporal coherence, i.e. phase synchrony differed significantly between electrode pairs. In condition of face perception (but not in no perception condition) there were successively evidenced an increase in phase coherence between parietal and occipital areas coinciding with successful recognition, a sharp desynchronization between all recording sites correlated to the transition from perception to movement, and a final increase in phase coherence between frontal and temporal areas tagging the motor response. From these data the authors inferred the existence of "an integrative mechanism that may bring a widely distributed set of neurons together into a coherent ensemble that underlies a cognitive act", a mechanism clearly disassociated from power emission in brain tissue but no less endowed with functional significance, since both synchronization and desynchronization of neuronal assembly activities corresponded to behavioral conditions: synchronization to perception and movement, desynchronization to the transition from one cognitive state to the other. They were witnessing, for the first time in human brain research, that the mere temporal profile of a large scale, non-localized brain dynamism might subtend high level cognitive functions. Aren't we tempted to say, pushing to the end the same line of thought: coherence on behavioral level means coherence on the level of neuronal group formation, all else being equal at lower levels (single cells, synapses, molecules, genes, etc.)? It might be risky to interpret it in the sense of a causal autonomy of higher global with respect to underlying local levels, but if we mean descriptive autonomy, that is exactly what the authors argue in their attempt to formulate a theory of emergence accounting for their data: "the relevant variable required to describe these assemblies is not so much the individual activity of the components of the system but the dynamic nature of the links between them". Hence they conclude:

> Under this vision, the brain appears as a resourceful complex system that satisfies simultaneously the exogenous and endogenous constraints that arise at each moment by transiently settling in a globally consistent state. These novel views on the brain might throw light on the emergent principles that link neuron and mind, as the large-scale integration of brain activity can be considered as the basis for the unity of mind familiar to us in everyday experience (Varela et al. 2001: 237).

5. In-brain and between-brain coherence

Neuroscience has an object: the brain, not to say an isolated brain. Even brain imagery—a so-called direct view on the brain in act—only describes the brain's functional architecture in terms of statistic averages and baseline subtraction that mask between-subjects differences. Perception and action are still overwhelmingly represented today as internal processes in an isolated brain despite the fact of their retrieval from the flux of interactions of the subject with environment and other subjects. However, if intracerebral synchrony signals interaction (any communication of information) between neural assemblies in a brain, it is tempting to infer by analogy that between-brain synchrony might signal interaction between subjects engaged in communication. Traditional neuroscience used to deal with EEG and fMRI scanner—new social neuroscience deals with dual EEG and Hyperscan. Can we be satisfied to cheer at that progress without considering its uncritically accepted presuppositions? Speaking of communication to describe the interactions between brain regions, we knowingly use a metaphor. No objection to that. This is no longer the case when one inquires into the brain correlates of communication between human subjects engaged in a dialogic relationship. While extending the application of the paradigm of communication by neuronal coherence to the case of ordinary conversation one leaves a clearly metaphorical usage to switch to another mode of expression that deceptively looks like the literal use. The confusion that threatens us is between two very different language games: dialogic and diagnostic. On the one hand, the events in a conversation are narrative episodes of a biography or intentionally aimed at targets of designations and qualifications in the public space for speaking subjects to meet on a common ground and understand each other. On the other hand, one is only dealing with correlative activation foci appearing on the screen of brain imaging scanner. No explanatory gain is provided by the inadvertent commuting of meaning content to and fro: from the domain of conversational interaction between full persons to the domain of neuronal interaction between brain regions, and return. The appearance to the contrary is due to the fact that the tortuous path followed by the explanation makes one lose sight of its illusiveness. In the first place, talking about coherence in terms of communication facilitates the cognitive interpretation of Hebb's law for the formation of neuronal assemblies, which only says, referring to individual cells: "what fires together binds together". Conversely, the principle of effectiveness of synaptic connections is called upon to explain the difference between successful and unsuccessful communication. A move that is not without paying tribute to the vulgar prejudice concerning a necessary involvement of some internal state expressed by the sender of the message and eventually imported in the mind of the receiver, so that to communicate tends to be mistaken as "to communicate a mental state to the partner (Schippers et al. 2010)". Presumably, the dialogue of lovers enhances the feelings of both—nothing like that when filling up a form: regardless of the difference, it does not prevent a successful communication in both conditions. Hebb's law may well account for the effectiveness of the connection by synapses that work rather than those who do not work. Yet projecting the talk of “effectiveness of communication” upon the context of conversation and trying to cope with that effectiveness cannot but create the phantasm of a mental condition specific of successful communications:

> if the neural coupling across brains serves as a mechanism by which the speaker and listener converge on the same linguistic act, the extent of coupling between a pair of conversers should predict the success of communication… The findings shown here indicate that during successful communication, speakers’ and listeners’ brains exhibit joint, temporally coupled, response patterns (Stephens et al. 2010).

Extrapolating towards future technological conditions some speculate on Hyperscan: specifically, two subjects (S1, S2) in their respective scanner will interact by visual signals. S1 sees red or green screen. S1 sends S2 a signal (possibly misleading). S2 wins if he guesses the colour, otherwise S1 has won. Prognosticated result: a global phase coherence of respective brain oscillations including an activity focus in supplementary motor area (SMA) presumably more important in brain of sender than of receiver, etc. Let’s part company with such science fiction and pay attention to what the researchers are ultimately aiming at. According to their ingenuous statement, in the future each player will access and influence the other’s brain activity: such short-circuit of any behaviour in the world will reduce society to a correlation “between socially engaged brains” (Montague et al., 2002). Shall the philosopher be accused of bringing an unfair trial against neuroscience if he recalls that society is composed of persons, not of parts of persons, hence not of brains? Even if we do not take account of hazardous speculations of some scientists, it's pretty clear from that example that naturalization of phenomenology guided by the coherence metaphor is engaged in a dead end. But in the very field of neuroscience this window of communication with phenomenology seems to have already closed. Let’s see how.

6. A challenge to pheno-physical reductionism?

Getting rid of my possibly biased stage-setting, one must concede that the phenomena of brain synchrony and more generally the new path of research on the temporal dimension of brain activity might imply—this is what Varela suggested—a challenge to classical pheno-physical reductionism: the attempt at explaining out consciousness and cognition in terms of brain physicochemistry—but a challenge which is neither unanswerable nor unanswered. In order to measure the full extent of the challenge neuroscientists should not have let go as they did of Varela’s autonomy thesis. Certainly, Varela himself never claimed a fully-fledged autonomy for the autopoietic emergence of neuronal assemblies, in the sense of an independence from causal influences at lower levels of organization of the material substrate of the brain:

> The entire physicochemical constitution is in constant flux; the pattern remains, and only through the organizational invariance can the flux of realizing components be ascertained. In particular, although autopoietic systems are most certainly dissipative chemical systems, a purely matter–energy characterization misses entirely the specific architecture or material circuitry of these very specific chemical systems, which inaugurate the domain of autonomous individuals, and thus of life altogether (Varela, 1997).

However, he no less emphasized the non-locality and the lack of energy demand characterizing the emergence (and disappearance) of synchrony patterns in neuronal networks, a feature that makes these patterns good candidates to be recognized as cerebral correlates of consciousness. That allowed him to support a descriptive (if not causal) autonomy for the high level of description of brain metabolism in relation to the underlying causalities. In that particular he was hardly followed, except perhaps by a few who think they have found the Gestalt laws of organization in the conditions for the emergence of neuronal assemblies (Singer 1999). But the mainstream of researchers thought otherwise. As there is no more absolute synchrony in the brain than absolute simultaneity in a relativistic universe, the emergence of cell assemblies needed to be inserted in a web of entangled functional loops recruiting in different ways all pathways and relay stations of brain activity. (For a synopsis of the cortico-sub-cortical circuits which presumably subtend actions and emotions, see Rolls: *The Brain and Emotion* 1999). Contextualization in those loops is of a nature to definitely relativize the apparent autonomy of cell assembly dynamics in the cortex at the summit of descending influences from the centre to periphery (as exemplified in attention or in intention). Specifically, the fast rhythms of cell assemblies, presumed neural correlates of cognitive functions, are surrounded—and cannot but be influenced—both by faster rhythms of synapse and cell cycles on one side, and by slower rhythms like circadian, developmental, aging..., on the other. As it is unclear how the phenomenological field of experience—characterized as the retentio-protentional extended window of the "Now" of consciousness—emerges in such polyrhythmic entanglement, we are driven to the conclusion that the gap between physiological mechanisms and experience as it is subjectively lived has not yet been crossed.

In some quarters of recent neuroscience the lure of population (not single cells) coding through holistic activity network patterns seems to give way to a return to, or a refurbishing of Barlow's hierarchical view of brain functioning and to the primacy in that hierarchy of the role of single cell coding:

> Nerve cells are the only means we know about whereby items of information occurring in different part of the brain can be combined; sensory discriminations require the combination of information from different parts of the brain; therefore this operation must be performed by a cell, and if one could record from the cell that did this, one would obtain results at least as good as those of the whole animal. (Barlow 1985: 133-134)

Starting from the assumption that synchrony of activation of individual neurons determines a transitory assembly of neurons in brain (Hebb 1949), but remembering at the same time Barlow's lesson researchers try relating assembly synchrony to a target neuron. In lack of absolute simultaneity in brain circuits the synchrony of source neurons upstream must be relative to a target neuron that integrates the resultant of the sum of their activities downstream. So that solely the neurons discharging in the critical time window of a target neuron (10-30 msec) are synchronous and form a transitory assembly possibly endowed with cognitive significance. A neuron integrating a first assembly might pertain to a second assembly whose neurons activations were synchronous from the viewpoint of another neuron (possibly the member of a third assembly). Adding to that that the same neuron might pertain alternatively or concurrently to different assemblies without change in energy consumption: how such ceaselessly reshuffled groupings might possibly be available to the brain's higher level of representation—not to say to the conscious subject? (Buzsáki 2010)

According to the privileged neuronal coherence hypothesis, synchronization of oscillations as a normal mechanism permits the communication between neurons or neuron assemblies subtending perception and motor behaviour. Desynchronization of oscillations is also a normal process of decoupling neuron assemblies so as to turn possible the transition from one cognitive state or behaviour to another in experience. But synchronization and desynchronization need not be always functional. Associate to pathologic symptoms of neuromuscular diseases (tremor in Parkinson, Tourette syndrome, dystonia, etc.) is an abnormal unselective pattern of synchronization between motor system areas. A review of works on essential tremor in Parkinson disease shows impressive average maps of long distance non selective synchronization between controlateral primary motor cortex where oscillations are recorded at double the frequency of muscular tremor, and various areas whose oscillations are coherent with that motor cortex, including lateral and medial premotor cortices, somatosensory cortex, thalamus, basal ganglia and ipsilateral cerebellum (Schnitzler & Gross 2005). Such perturbation proves brain selective synchronization to depend on a subtle balance between excitatory and inhibitory pathways in the circuitry connecting basal ganglia to thalamus to cortex. Reciprocally, dementia in Alzheimer disease has been linked to a reduced synchronization of high frequency oscillations. The ideas of emergence and phase locking backing the coherence hypothesis with its neighbouring suggestion of the possibility of stabilization of cognitively functional activity patterns in the whole brain dynamics prepare us insufficiently to deal with dysfunctions.

Not to extend the list, we'll just add that a rebuff to the non-localization tendency of coherency theories is perceptible in the search for a controlling centre in the brain responsible for modulating or even inducing the rhythms recorded at the higher level of cortical neuronal assemblies, as if the need for explanation could not satisfy itself with the sole description of emerging spatiotemporal patterns of activity. In this regard, we cannot ignore the contributions of those who try answering the question of the sources of brain rhythms and eventually point at thalamus as a possible multi-purpose pace-maker for oscillatory activities anywhere in cortical areas. For example, subjects bearers of electrodes both on the scalp and in thalamus are submitted to a semantic memory retrieval task such as emitting a word in response to the presentation of a couple of related or unrelated words ("desert" + "hump" => "camel"). The result is that during semantic recall a drop in low-frequency rhythms power at 1-2s post-stimulus followed by an increase at 2-3s in fast rhythm power registered at thalamic and cortical electrodes. Advancing their theoretical model of thalamus as pacemaker for cortical rhythms, the authors emphasize the fact that here again all is a question of balance between excitatory and inhibitory influences from underneath, not of emerging stable patterns. Low-frequency rhythms controlled by inhibitory projections from thalamic reticular formation to thalamo-cortical cells counterbalance high-frequency rhythms controlled by excitatory cortico-thalamo-cortical pathway (Slotnick et al. 2002).

Conclusion

—Has phenomenology been satisfactorily naturalized in a neuroscience? It might well have been the case if the *autopoietic* interpretation of emergent network coherence in cortex had prevailed. But in a multi-scale neuroscience the intra-brain and inter-brain distant interactions at the basis of cognitive states enjoy no privilege in relation to the cascading levels of organization of the living organism, from molecules to behaviour to phylogenesis. The possible window of dialogue Varela opened up—if our interpretation is correct—between the neurophysiology of neuronal coherence and a constitutive phenomenology based on *Einstimmigkeit der Erfahrung* seem to have had a short life time indeed, since the current orientation of the research community rather favors embedding the time structure of transitory occurring neuronal assemblies in the larger context of brain rhythms at the multiple levels of organization of the living being. So, if one wanted an explanation in mechanistic terms of our capability of giving sense of being to salient episodes of experience provided they kept a seamless tuning, that question would be broken up in so many subquestions that their addition would leave one with a sense of indefinite postponement. A sobering experience, considering that the temporal structure of brain metabolism in the image that Varela displayed was of a kind to lure some phenomenologists in the belief of having reached a satisfactory response.

Note: Some would prefer to know more about the connection between neurophenomenology and *enactivism*, a concept that I did not use in this paper. To recap, neurophenomenology is a scientific program aimed at integrating the first person description of subjective experience and the third person explaining of cognition through brain dynamics. Enactivism, in cognitive science, is a stance alternative to representationalism and to dualism, conceives cognition as contingent upon sensori-motor interaction between an organism and its environment. A link between neurophenomenology and enactivism was provided to Varela and his followers by the phenomenology of the lived body developed by Merleau-Ponty in the wake of the later part of Husserl's philosophy.

Bibliography

Barlow, H.B. 1985. The role of single neurons in the psychology of perception. *The Quarterly Journal of Experimental Psychology*, 37A: 121-145.

Buzsáki, G. 2010. Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron*, 68: 362-385.

Hebb, D.O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York.

Hubel, D.H. Wiesel, T.N., 2004. *Brain and Visual Perception. The story of a twenty-five-year collaboration.* Oxford University Press, Oxford.

Husserl, E. 1929, *ms* C8II.

Husserl, E. 1935, *ms* BI16.

Montague, P.R., Berns, G.S., Cohen, J.D., McClure, S.M., Pagnoni, G., Dhamala, M., Wiest, M.C., Karpov, I., King, R.D., Apple, N., Fisher, R.E. 2002. Hyperscanning: Simultaneous fMRI during linked social interactions. *NeuroImage*, 16: 1159-1164.

Petitot, J., Varela, F.J., Pachoud, B., Roy, J-M. 1999. *Naturalizing Phenomenology. Issues in contemporary Phenomenology and Cognitive Science.* Stanford University Press, Stanford, California.

Rodriguez, E., George, N., Lachaux, J-Ph., Martinerie, J., Renault, B., Varela, F.J. 1999. Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397: 430-433.

Rolls, E. 1999. *The Brain and Emotion.* Oxford University Press, Oxford.

Schippers, M.B., Roebroeck, A., Renken, R., Nanetti, L., Keysers, Ch. 2010. Mapping the information flow from one brain to another during gestural communication. *PNAS,* 107, 20: 9388-9393.

Schnitzler, A., Gross, J. 2005. Normal and pathological oscillatory communication in the brain. *Nature Reviews Neuroscience,* 6: 285-296.

Singer, W. 1999. Neuronal Synchrony: A versatile code for the definition of relations? *Neuron,* 24: 49-65.

Slotnick, S.D., Moo, L., Kraut, M.A., Lesser, R.P., Hart Jr, J. 2002. Interactions between thalamic and cortical rhythms during semantic memory recall in human. *PNAS,* 99, 9: 6440-6443.

Stephens, G.J., Silbert, L.J., Hasson, U. 2010, Speaker-listener neural coupling underlies successful communication, *PNAS,* 107, 32: 14425-14430.

Varela, F.J. 1997. Patterns of Life: Intertwining Identity and Cognition, *Brain and Cognition* 34: 72-87.

Varela, F.J., Lachaux, J-Ph., Rodriguez, E., Martinerie, J. 2001. The Brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience,* 2: 220-239.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0011

Making enactivism even more embodied

Shaun Gallagher[*]
Department of Philosophy, University of Memphis;
School of Humanities, University of Hertfordshire;
School of Humanities and Social Inquiry, University of Wollongong
[*]*s.gallagher[]memphis.edu*

Matthew Bower
Department of Philosophy
University of Central Florida

Received January 2014; accepted August 2014; published Autumn 2014.

Abstract

The full scope of enactivist approaches to cognition includes not only a focus on sensory-motor contingencies and physical affordances for action, but also an emphasis on affective factors of embodiment and intersubjective affordances for social interaction. This strong conception of embodied cognition calls for a new way to think about the role of the brain in the larger system of brain-body-environment. We ask whether recent work on predictive coding offers a way to think about brain function in an enactive system, and we suggest that a positive answer is possible if we interpret predictive coding in a more enactive way, i.e., as involved in the organism's dynamic adjustments to its environment.

Keywords: enactivism; sensory-motor contingencies; affect; intersubjectivity; predictive coding.

Enactivism is one version of recently developed embodied approaches to cognition. It offers an approach that is more informed by phenomenology and pragmatism than other versions of embodied cognition, such as the extended mind hypothesis (Clark 2008; Clark and Chalmers 1998), and more radical than the kind of "moderate" (Goldman 2012) or "weak" (Alsmith and Vignemont 2012) embodied cognition found in theorists who locate the body "in the brain" or who translate all bodily action into body-formatted representations (e.g., Berlucci and Aglioti 2010; Gallese and Sinigaglia 2012; Goldman and Vignemont 2010). Enactivists have entered into various debates with these other approaches (see Di Paulo 2009; Gallagher 2011a; Thompson 2007), and clear lines have been drawn to distinguish the differences that involve questions about functionalism, the importance of the body, the way one is coupled to the environment, and so forth. At the same time, however, there are a number of differences to be found within the enactive camp itself, so that the "early" enactivism of Varela, Thompson and Rosch (1991), closely continued by De Jaegher and Di Paulo (2007) is not exactly the same as the "middle' enactivism of O'Regan and Noë (2001; Noë 2004), or the "latest enactivism" of Hutto and Myin (2013). There are some obvious differences among these authors simply in terms of pedigree: Varela et al. are strongly influenced by phenomenology (as well as Buddhism), O'Regan and Noë by cognitive science, and Hutto and Myin by analytic philosophy of mind. While enactivists should be pleased with these convergences of disciplines and traditions, we're going to suggest that they should also be critical with respect to the question about how embodied enactivism is.

To pose this critical question we want to focus on the work of those enactivists who emphasize the role played by sensorimotor contingencies in perception, especially O'Regan and Noë. While this is clearly an embodied approach to cognition, we want to argue that it is not as richly embodied as it should be. One starts to see this in some recent authors who have distinguished the O'Regan and Noë brand of enactivism as the 'sensorimotor approach' *in contrast to* the enactive approach (see, e.g., Kyselo and Di Paolo 2013; also Stapleton 2013). The focus on sensorimotor contingencies emphasizes a narrow conception of embodiment in terms of neuro-muscular function. In contrast to this narrow focus, we want to suggest that these versions of sensorimotor enactivism need to be made more embodied. We'll focus on three issues: (1) affectivity; (2) intersubjectivity; and (3) a dynamic attunement theory of brain function. Emphasis on these issues is consistent with the enactivism associated with Varela, Thompson, and Di Paolo, and offers a positive account of enactivism that goes beyond the incisive critique of representationalism offered by Hutto and Myin. We also note that although the sensory-motor contingency approach focused on perception, the enactivist project concerns more than just perception. That is, the project is to develop an account of cognition more generally, including perceptual judgments and higher-order cognitive processes such as deliberation, decision, memory, and so forth.

Our goal is to make clear that these issues are important ones for continuing development of the enactivist approach. Elsewhere we have developed some critical considerations with respect to the first two issues (Bower and Gallagher 2013). We'll briefly rehearse these considerations in the next two sections. We'll then focus on the third, and argue that to have a more embodied enactivism we need a different understanding of how the brain works. Specifically we want to explore recent predictive coding models from an enactivist perspective. In this regard we generalize suggestions about the "interactive brain hypothesis" made by Di Paolo and De Jaegher (2012).

Affectivity

An enactive account of cognition and perception integrates a variety of bodily factors into cognitive processes. The body, understood as what phenomenologists call the "lived body," includes the related notion of a "body schema" (Gallagher 2005). The role of the body schema pertains to motor control and precisely the kind of sensory-motor contingencies emphasized by O'Regan and Noë (2001) and Noë (2004); it facilitates interactions with one's surroundings, and it contrasts to the "body image," a term that designates the ways in which the body shows up for consciousness, in certain circumstances, as its intentional referent.

The lived body in its full sense, however, involves more than the sensorimotor body schema and body image. It involves the full ensemble of bodily factors that govern conscious life, but that operate in a pre-noetic fashion, below the level of conscious monitoring and manipulation. Such factors may or may not be accessible to conscious awareness. They include the large realm of affect. This rich affective set of bodily factors is partially constitutive of perception in ways that go beyond sensorimotor contingencies. An account that focuses only on sensorimotor contingencies falls short due to its neglect of the relevance of the affective aspects including proprioceptive and kinaesthetic aspects—factors that should be of high interest since they derive from movement and contribute to one's practical grasp of sensorimotor contingencies. Affective factors, however, also involve a complex motivational dimension that animates body-world interaction (Bower and Gallagher 2013; Stapleton 2013; Colombetti 2013).

Meaningful encounters with the world imply a perceiving agent with some basic motivation to perceptually engage her surroundings. Schemata of sensorimotor contingencies give an agent the *how* of perception, a tacit knowledge of potential sensorimotor engagements, without giving its *why*, which depends on latent valences that push or pull in one direction or another for attention and for potential sensory-motor engagement, reflecting, for example, a degree of desirability. "The endogenously originating motivational viscera of the body are just as important to perception as the exogenously oriented sensory-motor elements" (Bower and Gallagher 2013: 111).

Affects are not restricted to the domain of phenomenal consciousness, although they may certainly have an effect on what experience feels like. I may consciously experience the blues, or I may be unaware that my whole demeanor reflects the blues. Affect is deeply embodied even to the extent that affective phenomena may be constrained by the functioning of the circulatory system. For example, heartbeat influences how and whether fear-inducing stimuli (images of fearful faces, in the reported experiments) are processed (Garfinkel et al. 2013). When the heart contracts in its systole phase, fearful stimuli are more easily recognized, and they tend to be perceived as more fearful than when presented in its diastole phase. That is, the fact that, rather than brains in vats, we are flesh and blood creatures equipped with beating hearts, explains in part why we have just the sorts of affective states that we do.

Consider particular instances of the affects involved in hunger and fatigue. Somaesthetic factors such as hunger delimit our perception and action possibilities, as well as our cognitive possibilities. William James once noted that an apple appears larger and more invitingly red when one is hungry than when one is satiated. A recent study (Danziger et al. 2011) reinforced the idea that hunger can shape, and perhaps even distort, cognitive judgment processes. The study shows that the rational application of legal reasons does not sufficiently explain the decisions of judges. Whether the judge is hungry or satiated may play an important role.

> The percentage of favorable rulings drops gradually from ≈65% to nearly zero within each decision session [e.g., between breakfast and lunch] and returns abruptly to ≈65% after a [food] break. Our findings suggest that judicial rulings can be swayed by extraneous variables that should have no bearing on legal decisions. (Danziger et al. 2011: 1).

In one sense, such affective factors appear "extraneous" only if we try to think of cognition as something that is disembodied, although clearly they may be extraneous to the formal aspects of legal reasoning. In any case, it seems reasonable to think that this embodied affective aspect of hunger has an effect on the jurist's perception of the facts, as well as on the weighing of evidence, and doesn't appear out of nowhere just when the judicial decision is made.

Typically our embodied condition does not reflect a simple, isolated affect—rather, there is a cocktail, a mélange of aspects that make up one's affective state. After a day of trekking up a mountain, one's perception may be informed by a combination of hunger, pain, fatigue, troubled respiration, feelings of dirtiness, and the kinaesthetic difficulty involved in climbing. It's likely that that the mountain path looks more different and less challenging than after a good night's sleep, not because of certain objective qualities that belong to the path, but because of my affective state. Such affective aspects color my perception as they more generally constrain my being-in-the-world.

Affective phenomena are pervasively integrated into our perceptual and cognitive experiences (Pessoa 2013; Barrett and Bliss-Moreau 2009), as noted by many enactivists (Colombetti 2007; Ellis 2005; Thompson 2007; Thompson and Stapleton 2009). Shifts of attention may be led in one direction or another by the affective ebb and flow of what we experience. From a phenomenological perspective Husserl (2004) describes such affective states involving tension, resolution, exertion, unease, and satisfaction/dissatisfaction as modulating our perceptual (but not only perceptual) attention. Attention, in this sense, is embodied in a variety of related ways. In visual experience, for example, attending to something may involve squinting or opening the eyes widely, it may involve a contortion of the face all the way from the scalp down to a gaping mouth or pursed lips, and so on (Bergson 2001: 27-28).

Affect is closely related to the phenomenon of perceptual presence elaborated by Noë (2004), and specifically involves a sense of interest or investment. Perceptual presence is the sense one has of the perceptual accessibility of non-apparent aspects or sides of a perceived object, or, more broadly, of what is not directly sensed in the present moment (e.g., the side of the object that is not visible). The notion of "perceptual interest" (Bower and Gallagher 2013) denotes the affective sense of the stakes or the costs involved in exchanges with one's environment. This is not the same as Husserl's concept of the "I can," which signifies the intuitive possession of a sense of skill or competence. Rather, even if one is capable of accomplishing some feat in those terms, one might still not feel "up to the task," or not feel inclined to do the work it might take, which is the affective nuance the sense of interest is supposed to highlight. Thus, interwoven with that perceptual sense of presence is a sense of the affective stakes of making something available or present. To make something available involves definite costs in following through on transactions with environmental affordances. One's environment affords many possibilities for action, but each has its affective price tag, and they are not all equally affordable. One thus not only has a practical (sensorimotor) understanding of accessibility, but an affective take on that same accessibility, in terms of interest or inclination to follow through. The latter may also involve a perceptual sense of the ease or difficulty of making something present.

It's true that a perceiving agent's perceptual stance is determined by a mastery of sensorimotor contingencies needed to access environmental affordances in suitable ways. Such mastery, however, once acquired, may be a relative constant, and more or less generic or standard set of skills suitable for most transactions with the world. Anyone with that same skill set might perceive in the same way. In contrast, the particularities of affect will differ from one individual to another, or from one day to another. Taking affective phenomena into account importantly enriches one's understanding of perception, since it clarifies the nature of individual perspective in perception. A broad spectrum of individual life circumstances may, in terms of affect, be brought to bear on perception, as well as other forms of cognition. These circumstances include not only physical burdens and impediments, such as the impediment of fatigue from physical exertion, but also broader circumstances having to do with time of day, since one typically is energized at the start of the day and tired out toward the end, or with longer-term life phases, since youth and old age surely shape one's perceptual interest.

While everyone is affected by such circumstances in one way or another, each individual lives them out in a unique way. The sense of the presence of the other side of an object, of what is behind one, of what is in an adjoining room, and the like, touches—over and above one's generic strategies for bodily coping with the environment—one's individual condition with all of its strengths and weaknesses. In a very simple example, what Noë calls the 'grabbiness" of an object is dependent not only on one's sense of the pertinent sensorimotor contingencies, whether the object is near or far, and properly shaped and weighted, etc., and not only on whether one is in a state of pain, or fatigue, or fear, etc., but on whether one is even concerned about (or inclined to) the possibility of grabbing the object.

Intersubjectivity

A series of experiments conducted by Proffitt et al. (1995; Proffitt 2003) purportedly show that the estimation of distance is influenced by anticipated effort. Subjects saddled with a heavy backpack tend to overestimate perceived distance, whereas those without backpacks do not. Proffitt et al. (1995) similarly describe how subjects overestimate the degree of incline of a slope when fatigued, and this may translate into the subject's lack of inclination (to climb), which further informs perception. The hill looks not only steeper, but also uninviting. This might count as at least indirect evidence that perceptual experience and/or judgment is informed by one's present affective state. These results, however, have been challenged by Durgin et al. (2009), in a way that nicely points in a different, but equally important direction for our considerations here. They show that steeper estimates of incline while wearing a backpack "are judgmental biases that result from the social, not physical, demands of the experimental context" (p. 1). Without awareness of this bias, subjects who sense the aim of the experiment estimate a steeper incline than subjects who are misled about the purpose of the backpack (e.g., that it contains electromyographic equipment to measure muscle tension). In other words, those subjects who had a sense of the experimenters' intentions were biased in favor of those intentions, without necessarily knowing it.

If Proffitt is right (see Proffitt 2009; 2013 for further discussion), his experimental results could be taken to point to the embodied-affective nature of perception. But if Durgin et al. (2009; 2012) are right, their results still point to an embodied phenomenon—namely, the significance of others on our perceptions and/or judgments.[53]

[53] We note that there is continuing debate about whether these are effects on perception itself or on perceptual judgments, a distinction that can be clearly made in the experimental lab (see Firestone and Scholl 2014 experiments on the El Greco fallacy). This is an important issue in regard to claims about cognitive penetration of perception, but for our limited purposes here it is sufficient

There are several ways to understand intersubjectivity as an embodied phenomenon. For example, some simulation theorists, like Gallese (2001; Gallese and Sinigaglia 2012) understand empathic consciousness or social cognition to be embodied in at least a weak sense. That is, they understand social cognition to depend on body-formatted neural processes, specifically the activation of mirror neurons (MNs).[54] A more enactive approach, however, interprets the mirror system to be activated in preparation for or anticipation of a response to the other. That is, MN activation does not ordinarily involve matching (in one's own system) or imitating the action of the other person (see Catmur et al. (2007); Dinstein et al. 2008; Csibra 2005 for empirical evidence; also Gallagher 2008); it involves anticipatory processes that are keyed in with affordances for further interaction, or preparation for a complementary action in response to an observed action (Newman-Norlund et al. 2007). In other words, it is part of a response to social affordances.

On the enactive view, social cognition is characterized by, and sometimes constituted by embodied interaction (Di Jaegher, Di Paulo and Gallagher 2010). This view is usually worked out in contrast to theory of mind (ToM) approaches that emphasize mindreading by either theoretical inference or simulation (Gallagher 2001; 2005; 2008). Intersubjective interaction is not about mindreading the mental states of others, but about directly perceiving their intentions and emotions in their postures, movements, gestures, facial expressions, vocal intonations, etc., as well as in their highly contextualized (by physical environment, social roles, culture, etc.) actions (Gallagher and Varga 2013).

One important aspect of intersubjective interaction that shows up even in cases where we are not explicitly interacting with the other, although another person is present, is the effect that his or her presence has on learning and perception. One can understand this developmentally in regard to how we learn what is important. Studies of "natural pedagogy," where there is explicit interaction, show that how the caregiver relates to the child influences what the child learns. Natural pedagogy, which involves ostensively directing the infant's attention to some object or event,

> enables fast and efficient social learning of cognitively opaque cultural knowledge that would be hard to acquire relying on purely observational learning mechanisms alone.... [H]uman infants are prepared to be at the receptive side of natural pedagogy (i) by being sensitive to ostensive signals that indicate that they are being addressed by communication, (ii) by developing referential

that there are such effects on cognition, whether perception, perceptual judgment, memory, etc. and specifically in everyday pragmatic environments.

[54] For the notion of body-formatted (or B-formatted) representations, see Goldman and de Vignemont (2010), and Goldman (2012) who propose a "moderate" or "weak" (Alsmith and de Vignemont 2013) conception of embodied cognition.

> expectations in ostensive contexts and (iii) by being biased to interpret ostensive-referential communication as conveying information that is kind-relevant and generalizable. (Csibra and Gergely 2009: 148).

More generally, it is through our interactions with others that we learn what objects are significant or valuable. We learn to understand the world along these lines of significance and value, and often objects that fall outside of such lines don't even register. In the same way that expert training hones the perceptual system so that experts are able to perceive things that non-experts fail to perceive, in some sense, we all become experts in everyday life through our interactions with others.

Indeed, this intersubjective education of perception and judgment continues throughout life. Adult subjects presented with a face looking towards (or away from) an object evaluate the object as more (or less) likeable than those objects that don't receive much attention from others. If one adds an emotional expression to the face and one get's a stronger effect (Bayliss et al. 2006; 2007). Furthermore, seeing another person act with ease (or without ease) toward an object will influence observers' feelings about the object (Hayes et al. 2007).

Social interactions, social roles and groupings also have their influence on how one perceives the world. In a modification of Proffitt's scenario, imagine being exhausted, but the incline is a hill that you are climbing with friends or to meet a loved one (Schnall et al 2008). Or, again, think of the affective import in situations where one would be seen by others as not up to the task, negatively impacting one's image. In some social circumstances one may find a particular setting to be of more interest and more attention grabbing than if one were with a different group, or alone.

The embodied and enactive brain

Following the weak embodiment strategy, one might argue that all such effects of affectivity and intersubjectivity are ultimately processed in the brain, so that even the most enactive aspects of perception are reducible to brain processes. After all, even if one interprets MNs as part of an enactive system, they are neurons located in certain brain areas. Moreover, when we look at how the brain works, we need concepts like representation and inference to explain it, and these go against strong enactivist claims that tend to be anti-representationalist and dismissive of the idea of subpersonal inferences. On the weak embodiment view (e.g., Goldman 2012), claims about enactive perception, affectivity, and intersubjectivity can all fit neatly into orthodox internalist accounts.

The notion of an enactive system requires conceiving of the brain in a different way. In evolutionary terms, the brain does what it does and is the way it is, across some scale of variations, because it is part of a living body that has hands that can reach and grasp in certain limited ways, and because it has eyes structured to focus, and so on. The sensorimotor system is the way it is because of the kind of organism the human body is. In addition the organism has an autonomic and peripheral nervous systems, and not just a central system. It attains an upright posture, which, in evolutionary terms reshapes essential features, including the brain (Gallagher 2005), allowing the person to cope with specific kinds of environments, and with other people. Changes to any of the bodily, environmental, or intersubjective conditions elicit responses from the organism as a whole. On this view, rather than representing or computing information, the brain is better conceived as participating in the action.

The enactive interpretation is not simply a reinterpretation of what happens extra-neurally, out in the intersubjective world of action where we anticipate and respond to social affordances. An enactive interpretation of the MN system, for example, points beyond the orthodox explanation of information processing to the possibility of rethinking not just the neural correlates of perception or intersubjectivity, but the very notion of neural correlate, and how the brain itself works. More than this, it suggests a different way of conceiving brain function, specifically in nonrepresentational, integrative and dynamical terms (see, e.g., Gallagher et al. 2013; Hutto & Myin 2013).

This doesn't mean that we should simply turn the issue over to neuroscientists to adjudicate, or simply consult what the neuroscientists say about subpersonal neural processes. Most neuroscientists are Helmholtzian and would endorse the idea that the neural processes underlying perception are inferential and representational.[55] Even the Bayesian predictive coding account treats perception and object recognition as an inferential process. As Friston puts it, predictive coding is "now a widely accepted view of perception that can be traced back to Helmholtz's original writings on unconscious inference" (2012: 248; also see Clark 2013).[56]

It's obviously important to understand brain dynamics. More generally, we've known for a long time that anticipatory processes are hugely important for perception and action. In neuroscience we have the work of Berthoz (2000),

[55] The Helmholtzian idea that perception involves subpersonal inferences may or may not be correct (see Bennett and Hacker 2003; Orlandi 2012; Hutto and Myin 2013).

[56] The unconscious processes performed by the brain "are like inferences insofar as from the observed effect on our senses we arrive at some conception of the cause of this effect. This is the case even though we only in fact have direct access to the events at the nerves; we sense the effects and never the external objects" (Helmholtz 1867: 430).

for example; and in phenomenology, Husserl's account of temporal structure remains important (1991; Gallagher 1998; 2011b). Work in predictive coding provides an account of how neural processing participates in these pervasive dynamic anticipatory processes. But again there are questions of how to interpret what we are learning about such processes. The assumption about visual perception, for example, is that the brain has no direct access to the outside world, so it needs to interpret or decode neuronal firing patterns that are generated by light hitting our retinas. Sensory data "is all the brain has access to" (Hohwy 2013, 13). If you think of this in terms of inference then the brain is seemingly deducing to the best explanation of what has caused a particular pattern of neuronal activation and thereby representing that cause via a process of causal inference. Since a given pattern could be caused by any number of different stimulus configurations, the task involves figuring the probabilities based on current neural states of the system that may relate, for instance, to context.

How does this Bayesian process work in strict neuronal terms? There is general agreement that the process is a hierarchical one involving synaptic inhibition based on an empirical prior—something that depends on the organism's previous experience and context-sensitive learning. This means that specific neural networks, currently in a particular state because the organism has previously encountered a particular stimulus or environment, or has a particular history, determine ongoing processes that lead to top-down synaptic inhibition (a modulation of connections) anticipating further processing consistent with prior processing. Such inhibitory patterns constitute a prediction which is then matched against ongoing sensory input. If there is a mismatch, i.e., if the new stimulus generates a different firing pattern than the one anticipated, prediction errors are sent back up the line and the system adjusts dynamically back and forth until there is a relatively good fit. So on the predictive coding model, this or something like this is going on when we perceive the world.

If that's the case, if these are the kinds of things that are happening in the brain, it's not clear that we need to think of it as a kind of inference rather than a kind of dynamic adjustment process in which the brain, *as part of and along with the larger organism*, settles into the right kind of attunement with the environment—an environment that is physical but also social and cultural (Bruineberg and Rietveld 2014).

Whether or not we should think that beliefs, values, as well as affective states and perspectives can shape the way that one quite literally sees the world, or can enter into subpersonal processes in terms of predictive coding models (Stapleton 2013), it remains an open question about how the neural (synaptic-inhibitory) processes described by such models are best characterized—whether as inferential (e.g., Hohwy 2013), or as part of a dynamical attunement of organism to environment, and/or in terms of plasticity where sensory-motor neurons have become attuned by associative processes and prior experience. In any case, on the enactive view, the explanatory unit of perception (or cognition, or action, etc.) is not the brain, or even two (or more) brains in the case of social cognition, but a dynamic relation between organisms, which include brains, but also their own structural embodied features that enable specific perception-action loops involving social and physical environments, which in turn effect statistical regularities that shape the structure and function of the nervous system.

The question is, what do brains do as part of a dynamical attunement of organism to environment in the complex mix of transactions that involve moving, gesturing, and interacting with the expressive bodies of others, with their eyes and faces and hands and voices; bodies that are gendered and raced, and dressed to attract, or to work or play; bodies that incorporate artifacts, tools, and technologies, that are situated in various physical environments, and defined by diverse social roles and institutional practices?

Di Paolo and De Jaegher (2012) have proposed the Interactive Brain Hypothesis in order to help model the possible relations between social interaction and neural processes. "The hypothesis states that interactive experience and skills play enabling roles in both the development and current function of social brain mechanisms, even in cases where social understanding happens in the absence of immediate interaction" (p. 1). We can extend this idea beyond the intersubjective context by following the suggestion that the brain is primarily and more generally an organ of relation (Fuchs 2011). Evan Thompson (2014) provides a good analogy to indicate that the mind is relational, and that the brain plays its part in that relationality. Saying that cognition is in the brain is like saying that flight is inside the wings of a bird. But just as flight doesn't exist if there is just a wing, without the rest of the bird, and without an atmosphere to support the process, and without the precise mode of organism-environment coupling to make it possible, so cognition doesn't exist if there is just a brain without bodily and worldly factors. Can this way of thinking be made consistent with predictive coding models?

Barrett and Bar's *affective prediction hypothesis* suggests a positive answer. On their predictive coding model "responses signaling an object's salience, relevance or value do not occur as a separate step after the object is identified. Instead, affective responses support vision from the very moment that visual stimulation begins" (Barrett and Bar, 2009, p. 1325). At the earliest point of visual processing, the medial orbital frontal cortex is also activated initiating a train of muscular and hormonal changes throughout the body, "interoceptive sensations" from organs, muscles, and joints associated with prior experience, and integrated with current exteroceptive sensory information that helps to guide response and subsequent actions. In this respect, perceiving the environment involves not just a set of neurons firing in the brain, but also undergoing certain bodily affective changes that accompany this integrated processing. Before we consciously recognize an object or other person, for what it or he or she is, our bodies are already configured into overall peripheral and autonomic patterns based on prior associations. In terms of the predictive coding model used by Barrett and Bar, priors, that include affect, are not just in the brain, but involve a whole body adjustment to the environmental stimulus.

This suggests that the brain participates in a system, along with eyes and face and hands and voice, and so on. And the brain would work differently if its embodiment lacked eyes, face, hands, voice, and so on. This is a fully embodied system that enactively anticipates and responds to its environment. How an agent responds and what an agent perceives will depend to a great degree on the overall dynamical state of the brain, but also on environmental factors, embodied affective and intersubjective factors, the person(s) with whom she is interacting, her worldly and intentional circumstances, the bodily skills and habits she has formed, her physical condition, as well as her history of personal experiences, and what the other person may expect in terms of normative standards stemming from communal and institutional practices (Gallagher et al. 2013). Change any of these things and we can expect changes in neural processing, not because the brain represents such changes and responds to them in central command mode, but because the brain is part of the larger embodied system that is coping with its changing environment.

Conclusion

We've argued that it's not enough to model an enactivist approach to perception and cognition on sensorimotor contingencies alone, even if they do play an important role in such matters. There is good evidence that affective and intersubjective aspects of embodiment are also important contributories to perceptual and cognitive processes. In addition, this realization pushes us to re-think the role played by neuronal processes in the brain. Even a neuroscience that frames brain function in terms of predictive coding needs to recognize that the brain is part of a system that attunes to and responds to its environment in a way that enacts a meaning relative to the particularities of its embodiment.

References

Alsmith, A. J. T. and De Vignemont, F. 2012. Embodying the mind and representing the body. *Review of Philosophy and Psychology*. 3 (1): 1-13.

Barrett, L. F. and Bar, M. 2009. See it with feeling: affective predictions during object perception. *Philosophical Transactions of the Royal Society of London*. Series B, Biological Sciences, 364 (1521):1325–1334.

Barrett, L. F. and Bliss-Moreau, E. 2009. Affect as a psychological primitive. *Advances in Experimental Social Psychology* 41, pp. 167-218.

Bayliss, A.P., Paul, M.A., Cannon, P.R. & Tipper, S.P. 2006. Gaze cueing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review* 13: 1061–6.

Bayliss, A.P., Frischen, A., Fenske,M.J. & Tipper, S.P. 2007. Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition* 104: 644–53.

Bennett, M.R. and Hacker, P.M.S. 2003. *Philosophical Foundations of Neuroscience*. Oxford: Blackwell Publishing

Bergson, H. 2001. *Time and Free Will*. F. L. Pogson (Trans.), Mineola, NY: Dover Publications.

Berlucchi, G. and Aglioti, S. 2010. The body in the brain revisited. *Exp. Brain Res*. 200: 25–35.

Berthoz, A. 2000. *The Brain's Sense of Movement*. Cambridge: Harvard University Press.

Bower, M. and Gallagher, S. 2013. Bodily affectivity: Prenoetic elements in enactive perception. *Phenomenology and Mind* 2:108-131.

Bruineberg, J. and Rietveld, E. 2014. Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*. 8: 1-14

Catmur, C., V. Walsh, and C. Heyes. 2007. Sensorimotor learning configures the human mirror system. *Current Biology* 17(17): 1527–1531.

Clark, A. 2008. *Supersizing the Mind: Reflections on Embodiment, Action, and Cognitive Extension.* Oxford: Oxford University Press.

Clark, A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36(3), pp. 181-204.

Clark, A. And Chalmers, D. 1998. The extended mind. *Analysis* 58 (1): 7-19.

Colombetti, G. 2007. Enactive appraisal. *Phenomenology and the Cognitive Sciences* 6: 527-546.

Colombetti, G. 2013. *The Feeling Body: Affective Science Meets the Enactive Mind*, Cambridge, MA: MIT Press.

Csibra, G. 2005. Mirror neurons and action observation. Is simulation involved? ESF *Interdisciplines*. http://www.interdisciplines.org/mirror/papers/.

Csibra, G., and G. Gergely. 2009. Natural pedagogy. *Trends in Cognitive Sciences* 13: 148–153.

Danziger, S., J. Levav, L. Avnaim-Pesso. 2011. Extraneous factors in judicial decisions. PNA S.First published online: www.pnas.org/cgi/doi/10.1073/pnas.

De Jaegher, H., and Di Paolo, E. 2007. Participatory Sense-Making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences* 6, 485-507

De Jaegher, H., Di Paolo, E. and Gallagher, S. 2010. Does social interaction constitute social cognition? *Trends in Cognitive Sciences* 14 (10): 441-447.

Dinstein, I., C. Thomas, M. Behrmann, and D.J. Heeger. 2008. A mirror up to nature. *Current Biology* 18 (1): R13–R18.

Di Paolo, E. 2009. Extended life. *Topoi* 28 (1): 9-21.

Di Paolo, E., & De Jaegher, H. 2012. The interactive brain hypothesis. *Frontiers in human neuroscience*, *6*.

Durgin, F.H., Baird, J.A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. 2009. Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review* 16: 964–969.

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J., & Williams, M. 2012. The social psychology of perception experiments: Hills, backpacks, glucose, and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1582.

Ellis, R. 2005. *Curious Emotions: Roots of Consciousness and Personality in Motivated Action*. Philadelphia: John Benjamins Publishing Company.

Firestone, C., & Scholl, B. J. 2014. “Top-down” effects where none should be found the El Greco Fallacy in perception research. *Psychological Science* 25(1): 38-46.

Friston, K. 2012. Prediction, perception and agency. *International Journal of Psychophysiology 83* (2): 248-252.

Fuchs T. 2011. The brain: a mediating organ. *J. Conscious. Stud.* 18: 196–221

Gallagher, S. 1998. *The Inordinance of Time*. Evanston: Northwestern University Press.

Gallagher, S. 2001. The practice of mind: Theory, simulation, or primary interaction? *Journal of Consciousness Studies*, 8 (5–7): 83–107

Gallagher, S. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gallagher, S. 2008. Inference or interaction: Social cognition without precursors. *Philosophical Explorations* 11 (3): 163-73.

Gallagher, S. 2011a. Interpretations of embodied cognition. In W. Tschacher and C. Bergomi, eds. *The Implications of Embodiment: Cognition and Communication* (59-71) Exeter: Imprint Academic.

Gallagher, S. 2011b. Time in action. *Oxford Handbook on Time* (419-37). C. Callender, ed. Oxford: Oxford University Press.

Gallagher, S. 2007. Simulation trouble. *Social Neuroscience.* 2 (3-4), 353-65.

Gallagher, S. and Varga, S. 2013. Social constraints on the direct perception of emotions and intentions. *Topoi.* DOI 10.1007/s11245-013-9203-x

Gallagher, S., Hutto, D. D. Slaby, J. and Cole, J. 2013. The brain as part of an enactive system. *Behavioral and Brain Sciences* 36 (4): 421-422.

Gallese V. 2001. The 'shared manifold' hypothesis: From mirror neurons to empathy. Journal of Consciousness Studies 8: 33–50.

Gallese, V., & Sinigaglia, C. 2011. What is so special about embodied simulation? *Trends in cognitive sciences*, *15*(11), 512-519.

Garfinkel, S., Minati, L. and Critchley, H. 2013. Fear in your heart: Cardiac modulation of fear perception and fear intensity. Poster presented at the *British Neuroscience Association Festival of Neuroscience*. (8 April 2013).

Goldman, A. and De Vignemont, F. 2009. Is social cognition embodied? *Trends in Cognitive Sciences* 13 (4), 154-59.

Hayes, A. E., Paul, M. A., Beuger, B., & Tipper, S. P. 2008. Self produced and observed actions influence emotion: The roles of action fluency and eye gaze. *Psychological research*, *72*(4), 461-472.

Helmholtz, H. 1867. *Handbuch der Physiologishen Optik*. Leipzig: Leopold Voss.

Hohwy, J. 2013. *The Predictive Mind*. Oxford: Oxford University Press.

Husserl, E. 2004. *Wahrnehmung und Aufmerksamkeit: Texte aus dem Nachlass (1893-1912)*. T. Vongehr and R. Giuliani, eds. Dordrecht: Springer.

Husserl, E. 1991. *On the Phenomenology of the Consciousness of Internal Time (1983-1917)*. Trans. J. Brough. Dordrecht: Kluwer Academic Publishers.

Hutto, D. and E. Myin 2013. *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge, MA: MIT Press.

Kyselo, M. and Di Paolo, E. (in press). Locked-in syndrome: A challenge for embodied cognitive science. *Phenomenology and the Cognitive Sciences.*

Newman-Norlund, R.D., M.L. Noordzij, R.G.J. Meulenbroek, and H. Bekkering. 2007. Exploring the brain basis of joint attention: Co-ordination of actions, goals and intentions. Social Neuroscience 2(1): 48–65.

Noë, A. 2004. *Action in Perception*. Cambridge, MA: MIT Press.

O'Regan, K. and Noë, A. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 23, 939–973.

Orlandi, N. 2012. Embedded seeing-as: Multi-stable visual perception without interpretation. *Philosophical Psychology 25* (4): 555-573.

Pessoa, L. 2013. *The Cognitive-Emotional Brain: From Interactions to Integration*, Cambridge, MA: MIT Press.

Proffitt D. R. 2009. Affordances matter in geographical slant perception. *Psychonomic Bulletin & Review* 16: 970–972.

Proffitt, D. R. 2013. An embodied approach to perception by what units are visual perceptions scaled?. *Perspectives on Psychological Science*, *8*(4), 474-483.

Proffitt, D., M. Bhalla, R. Gossweiler, and J. Midgett (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review*. 2(4). 409-428.

Proffitt, D., J. Stefanucci, T. Banton, and W. Epstein (2003). The role of effort in perceiving distance. *Psychological Science* 14(2), 106-112.

Schnall, S., Harber, K. D., Stefanucci, J. K., & Proffitt, D. R. 2008. Social support and the perception of geographical slant. *Journal of Experimental Social Psychology 44*(5), 1246-1255.

Stapleton, M. 2013. Steps to a ‘properly embodied’ cognitive science. *Cognitive Systems Research* 22-23, pp. 1-11.

Thompson, E. 2014. The embodied mind: An interview with Evan Thompson. Fall 2014. http://www.tricycle.com/interview/embodied-mind

Thompson, E. 2007. *Mind in Life: Biology, Phenomenology and the Sciences of Mind,* Cambridge, MA: Harvard University Press.

Thompson, E. and Stapleton, M. 2009. Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi* 28: 23-30.

Varela, F. J., Thompson, E. and Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: MIT Press.

interviews

AVANT, ISSN: 2082-6710
avant.edu.pl/en
OPEN ACCESS

Patterns of research

Interview with Shaun Gallagher[57] (Part II[58])

by Przemyslaw Nowakowski & Witold Wachowski
The interview was realized in January 2014.

What are the most appropriate questions in cognitive science as well as in philosophy of mind, in the second decade of the 21st century?

I think that cognitive scientists have to question their own reigning assumptions about how the brain works. There is a lot of hard data to look at, and it's quite overwhelming sometimes to analyze it, but at certain points we need to come up for air and try to sort out what it all means. For me the important questions are about assumptions we make in interpreting the data. Do we frame our interpretations in terms of classical computationalism and representationalism, or connectionist models, or dynamic systems theory, or predictive coding, or some other model? This may look like I'm putting the entire focus on the brain, but I think the motivation for these questions comes from the more embodied approaches to cognitive science. If the body and environment play an essential role in cognitive and affective aspects of existence, then we may have to rethink our basic models of how the brain works. Do we really want to maintain the Helmholtzian idea that the brain works on the model of inference? When we look at the actual processes that are specified by the predictive coding view, for example, should we really think of them as inferential?

How would you evaluate the current boom in predictive concepts of knowledge (in particular those inspired by Friston's works)? It is an interesting issue for us especially in the context of neurocentrism on the one side of the spectrum and explicit references to Varela's early works made by the supporters of these (predictive) propositions on the other. Do you think modern cognitive science too neurocentric?

---

[57] Humboldt Foundation Anneliese Maier Research Fellow 2012-17,
Lillian and Morrie Moss Professor of Excellence
Department of Philosophy, University of Memphis
s.gallagher[]memphis.edu

[58] see *http://avant.edu.pl/wp-content/uploads/S-Gallagher-interview-I.pdf*

Obviously the brain is one important element in cognition, as well as in other life processes. On the one hand, if you're a neuroscientist, then you need to be neurocentric in some sense. It would be odd advice to tell neuroscientists they shouldn't focus on the nervous system so much. We want them to gain as much knowledge about the brain as they can. On the other hand, I don't think that all of the other scientists who focus on cognition—developmental psychologists, experimental psychologists, linguists, roboticists, anthropologists, as well as philosophers—are overly focused on the brain. In this sense I'm not worried about neurocentrism in the actual doing of science. But let me add two caveats. First, cognitive neuroscience in general may be a bit too central-centrist—studying primarily the central cortical processes as if all answers should be found there. I think autonomic and peripheral processes are also important. I really like István Aranyosi's recent book on *The Peripheral Mind* (OUP 2013) and I think it should be widely circulated in the cognitive neuroscientific community. Second, I'm more concerned about neurocentrism in the popular media rather than in science itself. People, like Jan Slaby and Suparna Choudhury, who work in the area of critical neuroscience, raise important concerns about how all the cool things we are learning about the brain are covered in the media and are shaping the way the general public is thinking about human nature.

Going back to predictive coding and Friston's work, it's obviously important to understand the real dynamics of how the brain works. We've known for a long time that anticipatory processes are hugely important for perception and action. Neuroscientists like Berthoz have emphasized anticipatory processes in motor control; and in phenomenology we have Husserl's account of temporality with an emphasis on protention. Work in predictive coding, and as I understand it, Friston's work in particular, is important for providing an account of how neural processing participates in this larger and pervasive dynamic. But again there are questions of how to interpret what we are learning about such processes. I've tried to address this in my paper with Matt Bower (in *Avant*). Friston, and others, like Jacob Hohwy in his recent book on *The Predictive Mind*, associate predictive coding with Helmholz's notion of unconscious inference. For visual perception, for example, it is assumed that since the brain has no direct access to the outside world it has to interpret or decode neuronal firing patterns that are generated by light hitting our retinas. Accordingly, the brain is seemingly inferring to the best explanation of what causes a particular pattern of neuronal activation. This involves figuring the probabilities based on current neural states of the system—a current state that is generated by prior experience. So how does this 'Bayesian' process work? It involves top-down synaptic inhibition based on an empirical prior—which means the organism's previous experience and context-sensitive learning. Predictions are then matched against ongoing sensory input. If there is a mismatch, prediction errors are sent back up the line and the system adjusts dynamically back and forth until there is a rela-tively good fit.

Again, the question is how should we interpret this. That is, if this is literally what is going on in the brain—if we are talking about synaptic inhibition based on prior experience, and a dynamic process that involve prediction error correction, then it's not clear to me why we should think of it as a kind of inference rather than a kind of dynamic adjustment process in which the brain, as part of and along with the larger organism, settles into the right kind of attunement with the environment. What's going on in the brain is one integrated part of the dynamics that are not just inside the brain; they are transactional across brain and body and environment.

What is your opinion about the nature and value of interdisciplinarity? Is this phenomenon a natural step in the development of knowledge?

I'm a big fan of interdisciplinary teamwork in science. In previous centuries maybe Descartes could be expert in philosophy and mathematics and physics; likewise, Newton in physics and theology, and Locke in medicine, law, and philosophy; Hume in philosophy, economics and history; Adam Smith in economics and philosophy. And as late at the 19th century James could teach physiology, and psychology, and philosophy. There may still be a few geniuses who can claim expertise in more than one field, but given the high degree of specialization now required in these different fields, the rest of us need to work with others, and it often takes an interdisciplinary team to work on some questions. This can be difficult too, since different disciplines have different vocabularies and think about problems in different ways. But I always find this kind of interdisciplinary work to be a learning experience—one which involves a tension between maintaining standards in one's own discipline while trying to be flexible enough to accept some different ways of thinking about the problem that others bring to it. All such difficulties aside, I think the promise of interdisciplinary research is that it can bring different perspectives to bear on a problem. A good example—over the past two years I was involved in a project to study experiences of awe and wonder had by astronauts during space flight. This had never been studied scientifically before. The project involved neuroscience, psychology, and phenomenology, as well as linguistic analysis and hermeneutics. In the end we were able to report on the phenomenology of such experiences, but also on brain processes as they were tied to very specific environments (in this case, simulated space flight), and on the religious and cultural background of the individual subjects who participated in the experiment. The interdisciplinary approach gave us a much fuller picture, in different dimensions, than would be possible if we did just the phenomenology, or the neuroscience, or the psychology.

In this regard, then, can you say more about the role of phenomenology in experimental research nowadays? Is the present situation of phenomenology, as some claim, a little disappointing? How important is its role in the context of neurophenomenology?

I'm not sure what you mean by 'disappointing'. I would say it's a little different—that the situation of phenomenology is different than it was at the beginning of the 20th century. Maybe if some people think that it should be just the same as what Husserl proposed, then they may be disappointed. We are literally 100 years beyond the publication of Husserl's *Ideas*. For some people phenomenology is just the historical bit of philosophy developed by Husserl, Heidegger, Sartre, Merleau-Ponty, and so on. But phenomenology would be dead if that is all it was. Even in that history, Merleau-Ponty was already a far distance from Husserl, and I'm not sure that there is just one thing called phenomenology. On my view it would be sad if we were not using and developing phenomenology further. Certainly the original phenomenolo-gists would be sad if what they developed was now just history. I'm optimistic that phenomenology is still relevant to both philosophy and science.

Neurophenomenology is one way that phenomenology remains relevant. The study that I was just describing is, I think, a good example of neurophenomenology. It involves a correlation between neuroscientific data and first-person phenomenological report—and more than that… (You can find a full account of the study at
http://www.chdr.cah.ucf.edu/spaceandspirituality/publications.php,
and some specifics about the methodology we used at
http://www.frontiersin.org/Journal/10.3389/fnhum.2013.00608/full)
In effect, this study would have been impossible without employing a phenomenological methodology.

Your "A pattern theory of self" may be seen as a proposal aimed to integrate the research on the self. From our perspective, your division between the self-referential and the self-specific processes is fashioned into a kind of dispute between Northoff and Legrand on the neuronal basis of the self. Do you think that your inquiries follow the direction of Northoff?

My idea was to try to provide a theoretical framework that could integrate research on self—I agree with that. In some regard I was trying to remain neutral with respect to how precisely anyone should think about the different aspects that constitute the pattern. In addressing Northoff's research I was suggesting that there was a way that it could find a place within this theoretical framework, without necessarily saying that I agree with the way that he characterizes the self. I think that Legrand (Legrand and Ruby 2009) offers a critique of the kind of research that Northoff has done. Legrand and Northoff might themselves characterize this as a dispute. But I was again trying to indicate that Legrand's work also could find a place within the concept of a pattern theory of self. So I was suggesting that one could map both Northoff's research and Legrand's research into the broader and relatively coherent theoretical landscape of research on the self, and then be able to say some things about how they differ, without denying that there may be ways in which they are simply pointing to different aspects of self. Generally, from this perspective, I think that if someone would claim that their particular research on self captures all we need to know about self (and I don't think either Legrand or Northoff do this) then they would clearly be wrong. The question then is whether we can make different research approaches and conclusions consistent. Likely not; but the pattern theory may provide a framework in which we can work out such inconsistencies without ignoring the possibility that what is at stake is something much more complex. That's why I wrote that one "benefit of the pattern theory of self is that we can more clearly understand various interpretations of self as compatible or commensurable instead of thinking them in opposition" (2013a). But I also said that I didn't think this would solve all philosophical problems about self.

I wouldn't say that this moves us in the direction of Northoff. Indeed, I indicate that on my own view I think that Legrand and Ruby are right—and I point to a different paper where that agreement is reinforced (Gallagher 2012). But in the 2013 paper I wasn't outlining my own view, or my own pattern theory of self—I was proposing the concept as a neutral framework. Of course, the specifics that I included may in fact reflect my own biases in how I think of self, but if there are more aspects, or if one thinks that I offered too many parts to the pattern, then one can make the argument. I suppose the only way to disagree with that concept of a pattern theory is to say that the self is not a pattern at all, but just one thing that may or may not be one of the aspects I included in the pattern. And there are other interesting questions that can be explored. For example, what status does a pattern have—is it something real or something that is perceiver-relative?—important questions explored by Dennett (1991) and Haugeland (1993).

Your research spans a variety of topics—from self and psychopathology to education. Which of these problems are in your opinion currently most important?

I think a lot of these topics are interrelated, and I'm not sure I can say that one is more important than the other. You didn't mention intersubjectivity, and I think that is something that ties these other issues together, although they also relate in other ways. I think that if you focus enough on one topic it leads to the others. An obvious example is the connection between self and psychopathology, something emphasized by phenomenologists like Josef Parnas in Copenhagen and Louis Sass at Rutgers. Many psychopathologies also involve problems with intersubjectivity. Self-identity is also shaped by intersubjective and cultural forces, and this would certainly include educational institutions.

What is your opinion about social science studies on distribution of cognitive activity between human and non-human agents. We mean works by Bruno Latour, Michael Lynch etc. They seem to be of interest for some cognitive scientists, from Edwin Hutchins to Lambros Malafouris.

Yes, I find them interesting too. I've been developing the idea of a socially extended mind, by which I mean that our cognitive processes are extended not only by technological means or the kind of hand-held devices that Clark and Chalmers, following Hutchins, talk about, or artifacts (Malafouris), or environments (Kim Sterelny), etc.—all of which I think are important and interesting—but also by large institutional practices (Gallagher 2013b). I've focused on the legal system—a cognitive institution that helps us to solve certain problems. But science itself, understood as a cognitive system, may be an even clearer example. Latour's work is certainly relevant in this regard. So I think understanding social structures and what institutions do to us, not only in terms of cognition and problem solving, sometimes enhancing our cognitive abilities, and sometimes distorting them or limiting them, but also as human agents who live in intersubjective and social arrangements. Institutions can also distort our human relations—so these are topics not only for cognitive science, but also for critical theory.

What are your most important non-academic interests?

My family—most important. I take a non-academic interest in them. My wife and two daughters are doing interesting non-academic things and I try to keep up with them. My mother and sister and her family live in Philadelphia, my hometown, and I don't visit them enough. Living in Memphis is very cool—the music is fantastic, so I've been getting more interested in that, which is a return to something I loved when I was younger. Beyond that, food, wine, the beach and travel.

References

Dennett, D. C. 1991. Real patterns. *The Journal of Philosophy*, *88*(1), 27-51.

Gallagher, S. 2013a. A pattern theory of self. *Frontiers in Human Neuroscience,* **7 (**443): 1-7.

Gallagher, S. 2013b. The socially extended mind. *Cognitive Systems Research*. 25-26: 4-12.

Gallagher, S. 2012b. First-person perspective and immunity to error through misidentification. In *Consciousness and Subjectivity*, eds S. Miguens and G. Preyer. Frankfurt: Philosophical Analysis Ontos Publisher: 187–214.

Haugeland, J. 1993. Pattern and being. B. Dahlbom, ed. *Dennett and His Critics*. Blackwell.

Legrand, D., & Ruby, P. 2009. What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychological review*, *116*(1): 252.

AVANT, ISSN: 2082-6710
avant.edu.pl/en

The functionalist's body
Interview with Robert D Rupert[59]

by Przemyslaw Nowakowski & Witold Wachowski
The interview was realized in August–October 2014

In the history of philosophy and science, we may find many works that have been forgotten, even though it would seem that they anticipated some of the ongoing research in a particular way. Do you have your favorite older works that you think deserve to be appreciated (especially in the context of embodied or embedded cognition)?

I think Ron McClamrock's *Existential Cognition: Computational Minds in the World* deserves more attention than it has gotten in the 4E literature. It was published in 1995 and, in my opinion, should be counted among the classics from the dawn of the current 4E movement.

As we know, you support the embodied mind theory, but you remain skeptical about the extended mind theory. What properties of the body of do you consider particularly important for shaping our cognition?

The bodily properties that influence cognition the most are, I think, structural properties of motor commands and of our sensory-based interactions with the world. Of particular importance are structural properties having their roots in the temporal, spatial, and more generally geometrical profiles of motor commands and routines and sensory-based bodily interaction with the environment.

I support a weak form of the embodied theory, according to which neural representations of, or neural simulations of, bodily activity play a central role in cognitive processing. I don't claim that *all* human cognitive processing is carried out entirely in a sensori-motor code, but I suspect that a surprising large amount is. To the extent that our cognitive processing does involve representations in an amodal format (a non-motoric code that is not specific to any

[59] Professor of Philosophy, Department of Philosophy, University of Colorado at Boulder.
Professorial Fellow, School of Philosophy, Psychology, and Language Sciences, The University of Edinburgh.
Robert.Rupert[]Colorado.edu

sensory modality), I think that sensori-motor interactions cause those amodal representational units to appear in the cognitive system; they cause what might have otherwise been little more than snippets of neural noise to take on functional roles in cognition, and this process also affects the processing profiles of the amodal concepts so acquired. This suggests one clear way in which the nonneural body shapes the mind, at least if we take "shape" to be is used in a standard causal sense.

In addition, bodily experience or activity causes the formation and refinement of bodily representations—involved in motor control, bodily contact with the world, correlated sensory experience—and these representations (or their neural realizers) affect cognitive processing both via their causal connection to the formation of amodal representations (or to the realizers of amodal representations) and by the way in which they contribute to the control of behavior alongside co-contributing amodal representations that are directed toward the same objects as are the co-active, co-contributing bodily representations.

I'm not convinced, though, that the nonneural body does any actual cognitive processing or is properly part of the cognitive system. For example, when arguing for a strong embodiment thesis, Andy Clark appeals to Susan Goldin-Meadow's work on the cognitive role of gestures, which seems to show that gesturing during problem-solving changes the nature of the cognitive process; on this basis, Clark argues that the gestures literally constitute part of the relevant cognitive processing. I'm just as inclined, however, to think that what enhances cognition in these cases is the efferent copy of the motor command to gesture, not the gesture itself. These are open empirical questions, though, and it might be that the hands (one of the most likely candidates) contribute extensively to actual cognitive processing.

Jerry Fodor said: "Who could doubt that the mind is embodied? And given that we are all clear that the mind is embodied, where does that get us? I mean, everybody knows the mind is embodied, unless you've got religious stuff or something, some metaphysical or ideological biases, but now, right, I agree, the mind's embodied, and now what do I do? Search me!"

Do you think that the fact of embodied cognition modifies our view of cognitive processes (as in Shapiro's replacement hypothesis), or when we describe cognition as embodied, we merely complement the classic cognitive research?

Yes, the embodied perspective substantially modifies our view of cognitive processes and, at the same time, it may still be that the embodied view merely complements classic cognitive research. Let me explain.

I think the embodied perspective modifies our view of cognitive processing in at least the three following ways:

First, with regard to historical interaction and theories of representation. Here the embodied perspective helps us to see how the nonsemantically individuated units of cognitive processing arise and how that history affects not just the processing profile of those units—which I emphasized in response to the preceding question—but the representational value of those units. For example, by understanding the conditions under which some bit of neural flux becomes a functional cognitive unit, we see simultaneously the sensorimotor interactions that establish the representational content of that unit, that is, we locate the property, kind, or individual in the environment (typically) that the unit represents. The nonsemantically individuated unit represents the thing (broadly speaking) interaction with which led to the firming up of that unit as a cognitively relevant unit.

Second, with regard to the processing profile of cognitive units. For any representing unit that comes to play a role in cognitive processing, there are conditions under which that happens, and those conditions affect what that unit ends up doing. Moreover, these conditions normally have a bodily component, the nature of which can then contribute to the processing profile of the unit the use of which is thereby reinforced. These might be conditions that reinforce certain movements over others or cause certain kinds of experiences or thoughts to be more closely associated than they would have been.

Third, with regard to architecture. The embodied literature draws our attention to a wealth of results that implicate bodily processes in cognition. I think these results are best interpreted as part of a larger body of evidence, emerging from all quarters in cognitive science; this body of evidence suggests something about the human cognitive architecture, that it contains a massive number of representations, many families of which co-represent (they represent the same object, individual, or property), and that much of the variation in behavior we associate with conscious attention, or lack of it, is to be explained in terms of variation in the sheer number of co-representing units active. (Imagine lots of sticks being pushed from slightly different angles in an effort to get a huge stone to move; the more contributing sticks, the higher the chance the stone will go where it's "meant" to go. Similarly with mental representations: the sticks correspond to co-referring representations, and the moving of the huge rock is the body's skillful or attentive interaction with the environment.) The embodiment-related data seem to show that what we think of as distinctively bodily stimuli (including actually induced movement) increase the number of active neural representations in bodily formats or in somatosensory cortex. This results in a stronger contribution to behavioral control from those areas, but in the typical case, these units work in conjunction with co-referring representations not in a bodily code in order to control behavior.

At the same time, I sympathize deeply with the sentiment behind Fodor's remarks, partly because none of the contributions listed above go so far as the replacement hypothesis. Even when modelers and experimentalists focus on the interface between the organism and the environment—and here we might begin to think that bodily-action-in-the-world replaces the need for computational processing—the typical cognitive model remains of the same sort as traditional computational and representational ones, at least with regard to its abstract characteristics of the sort Fodor has historically cared about.

Here's what I have in mind. So far as I can tell, most proponents of strong embodiment theses misunderstand the metaphysical relation of their views to classical cognitive research; as a result, they claim that their strongly embodied views stand at odds with functionalism and computationalism (and they typically disparage these classical views as "disembodied"). This seems to me to be simply wrong. According to the classical view, the physical body (the brain, in particular, but this applies to whatever matter realizes the computational system) determines, in the strict metaphysical sense, which functions the human cognitive system computes. That is the very nature of the realization-relation. Thus, on the classical view, there's a one-way determination relation running from body to mind; the body is very deeply "in charge" and is the root of our cognitive beings. Therefore, unless we construe embodied theories as type-type identity theories—which raises a host of problems—embodied views are a complement to orthodox views. Embodied approaches do make an important contributions; they entail that human computational (or otherwise functional) cognitive processes are best described in a fair degree of detail that, practically speaking, can be discovered only by examination of bodily structures and the ways in which humans interact with their environments. But, so long as the embodied theorist isn't offering type-type reduction (the identification of types of cognitive or mental states or processes with bodily types of states and processes), the embodied view is no less functionalist, and no less disembodied, than the orthodox view, at least metaphysically speaking.

Granted, epistemologically speaking, the embodied approach recommends a different methodology than was pursued by many classical computationally oriented cognitive scientists. The embodied functionalist lets bodily activity be her guide, epistemically, when attempting to figure out which algorithms govern human cognition and where they're realized (and how many shortcuts the embodied computing system can take given stabilities in the environment). But, this has nothing to do with the truth of functionalism but rather it stands in opposition to a certain empirical bet that many computationalist-functionalists made in the early days of cognitive science: that the relevant algorithms and the location of the machinery that executes them could be identified from the armchair.

Although there is a lot of research on embodied cognition, some authors still explore the relationship between these issues and the classical epistemological issues. Do you think that we should have a specific embodied epistemology? Or would that be a completely unnecessary effort, as we already have conceptions compatible with the notion of embodiment, and if so, which concepts may be used in this way?

I do think a situated perspective has some important implications for traditional epistemology, partly by bolstering the general case for externalism about justification (or warrant). But, to my mind, one of the most interesting ways in which the embodied perspective enters into epistemological debates is via its bearing on philosophical methodology, that is, on the epistemology of philosophy itself. Timothy Williamson has argued that the supposedly exhaustive dichotomy "*a priori* versus *a posteriori* justification" is sometimes inapt, that its application doesn't shed light on philosophical matters of great interest; such matters involve possession and application of philosophical concepts (such as the concept of justification itself) that cannot be reduced to a sensory base and at the same time do not ground analytic truths known *a priori* by all who grasp these concepts. Williamson emphasizes the extent to which bodily experiences shape our intuitions about concept application (to hypothetical cases, for instance) and argues that two individuals who share the same concept might have had their intuitions shaped by different series of interactions with the environment (and the differences in these interactions is not a matter of having different stored sensory impressions, in the traditional empiricist's sense). Perhaps without intending to, it seems to me that Williamson presents an embodied conception of the formation of philosophical intuition and the way it guides responses to thought experiments. People who share the same concepts, by many standard epistemological or semantic measures, have differing embodied experiences connected to that concept. As a result they make conflicting judgements about, or have contrasting intuitions in response to, philosophical thought experiments. These embodied experiences don't constitute part of a "reduction base" of experiences out of which concepts are built in empiricist fashion, but neither are the judgments they lead subjects to make constitutive of the concepts in question and given *a priori*. I think Williamson's exactly right (or at least, I think he is on my gloss of his work!).

Another important epistemological debate concerns the status of the embodied approach itself. On some more extreme versions of the embodied view (Lakoff's, for example), the question of truth doesn't arise. Folks who eschew truth and objective reality (and think instead that we enact our own worlds) should do more to explain the epistemological status of their own views or to develop a view of cognitive processing that gives an epistemological (and corresponding metaphysical) anchor to the embodied view itself. I would find the embodied view uninteresting if its strongest advocates could at best say "we believe in the embodied view because the fine-grained details of our body made us do it"! So, I think there's pressure on philosophers to develop an embodied epistemology (and accompanying metaphysics) that underwrites their own theorizing and theoretical conclusions.

Which contemporary experimental works, in your opinion, are changing or will change our view of embodied (or embedded) cognition?

I think that much of the work on the neural basis of decision-making, planning, and abstract reasoning is of special importance, with regard to embodied cognition. It's helping us to gauge the extent, for example, of visceral contributions to these supposedly purely rational processes. We have to go beyond the mere observation of activation of motor cortex during so-called higher cognitive processing and discover the role (say, via articulatory encoding) that motor processes play in higher cognition.

Two threads of contemporary cognitive neuroscience seem especially interesting to me. The first involves the exploration of neural re-use and redeployment, the idea that what have been thought of as various functionally specific areas of the brain each participates in a fairly wide variety of forms of cognitive processing. The second involves decoding, the use of machine learning resources to extract signatures distinctive of different mental or cognitive processes (of, say, physical pain versus emotional pain, or of differing visual images). A combination of these kinds of experimental work has the potential to offer substantive embodiment-related insights. I'm inclined toward a hybrid view—to think that embodied representations and simulation make up only part of our cognitive resources, but resources that are constantly in use and contributing to the full range of cognitive processes. So, in my ideal world, decoding is used to identify the alphabet of neural computing (to identify the "workings" as Michael Anderson calls them), which I suspect will reveal that many of these fundamental components are in fact deeply connected to bodily experience and contribute their bodily content to a wide range of cognitive processes.

It seems that there are several versions of enactivism, or at least several approaches known as enactivism. Some advocates of enactivism emphasize the convergence, or even the conflation, of cognition and life[60]. Others indicate that there are elements of constructionism in this approach. Still others see enactivism as based on law-like relations called sen-

[60] By the way, it is worth mentioning the words of G. Bateson: from an enactivist perspective "epistemology and theories of mind and theories of evolution are very close to being the same thing" [Bateson, G. 1987. Men are grass: Metaphor and the world of mental process. W.I. Thompson, ed. *Gaia, a Way of Knowing: Political Implications of the New Biology*. Hudson, NY: Lindisfarne Press: 37-47].

sorimotor contingencies. Could you describe what valuable insights may be prompted by these perspectives in the studies on cognition?

In my opinion, the most instructive versions of enactivism emphasize the contingency of the range and kinds of properties in the environment that an individual human comes to interact with, exploit, and represent. It serves no good purpose, in my opinion, for the enactivist to claim that the subject's acting on the environment „brings a world into existence." In order for the interaction to take place, the world must already be there! But, the enactivist is right about the following aspect of cognition: which of the many, many properties in the world the subject represents or becomes sensitive to depends on the subject's previous interactions and her biologically given capacities and needs. There's a kind of canalization that takes place—over the course of individual development, over the course of an entire life (say, in the gradual acquisition of expertise), and over the course of smaller-scale cognitive interactions (in which one kind of output from the agent renders salient to the agent a different aspect of the environment than would have been salient to her had she taken a slightly different action, with a „snowballing" effect).

"Alternative" approaches in cognitive sciences are undergoing a sort of a conceptual turmoil. Some authors propose a notion of "3E Cognition" or "4E Cognition" (Embodied, Embedded, Extended, Enacted) as a kind of a core, or a sheltering umbrella. Are your arguments—included, among others, in *Cognitive Systems and the Extended Mind*—insufficient?

In chapter 1 of the book, I challenge the utility of that umbrella. The various approaches are genuinely distinct; and in some ways, they are in tension with each other while in many other ways they are „mix and match." For example, an extended cognitive system might or might not proceed with its cognitive tasks in a heavily embedded way. Having disinguished these various views, I proceed, in the first two major sections of the book, to raise various objections to the extended view in particular and to the arguments in support of it, especially when it's offered as a new paradigm in cognitive science or the basis of a scientific revolution. But, I also offered a positive theory of cognition and cognitive systems (in the book and in more recent papers), one that grounds various objections to the extended view (and of arguments in support of it) but that also allows, at least in principle, for the appearance of extended cognition. I think that's still where the situation stands.

I'm a staunch supporter of the embedded view, and combined with my view of cognition as the activity of a persisting collection of integrated mechanisms, it seems to entail that human cognition takes place inside the organism but depends on and exploits external (noncognitive) resources to a surprising extent. I'm quite impressed by much of the same empirical work proponents of the extended mind are impressed by (Ballard's work on deictic pointers, for instance); in my opinion, though, most of that work supports the organismically oriented embedded view, as opposed to the extended one.

I'm also quite committed to a weak embodied view of the human cognitive system. Many interpreters of body-related experimental work claim that it reveals the bodily constituents of cognition or that it shows that linguistic meaning is inherently embodied. I draw a different conclusion. To my mind, the experimental results on embodiment show that the cognitive system contains an enormous number of representations that are being activated in parallel, and many of them represent the same thing (take the same object, or corefer); moreover many of these appear in body-related areas of cortex or represent bodily processes, and their activation alongside amodal but co-representing units affects cognitive processing in measurable ways.

I think the lumping of all of the E-views together has the potential to cause a lot of confusion; after all, some of the most well-known enactivists reject representation, yet many proponents of other 4E views embrace representations. So, it might seem like these various approaches share too little to constitute a coherent view. Then again, some of the basic conceptual ideas behind many of these approaches may provide a dovetailing vision of cognition. And here I mean to go beyond just the idea that 4Es have a common enemy, a certain kind of traditional cognitive science (perhaps a straw man, in actuality). Rather, I'm claiming that, if wedded together, one might get a picture of human cognition that looks very different from what one might have extracted from such early models of cognition as the General Problem Solver. Here's the possibility I have in mind: There is an organismically centered cognitive system that relies heavily on simulations of bodily processes to guide the active exploitation of resources in the environment, most often taking up those resources in a fleeting way that does not render those resources truly cognitive, but sometimes in durable ways that add new external components (TVSS perhaps?) to the otherwise bodily bounded cognitive system.

How do you think, are you rather seen as a critic, or rather as an advocate of any approach?

I would guess that more philosophers know me as a critic of the extended view (and the view that groups of humans have minds or cognitive systems) than as anything else. But, I think some of my other positive work is reasonably well known, for example, the Best Test Theory of representational content, the component-forces theory of ceteris paribus laws, and my developmental views about the origins of representations. And, it's worth noting that some distinctively positive views have come out of my critical work on extended cognition, including a theory of cognition (that it is constituted only by activity in a collection of mechanisms that contribute in overlapping subsets to the production of a wide range of forms of intelligent behavior). Grappling with embodied views has also led to the formulation of my view that the mind is massively representational and wholly subpersonal (in the sense that there is no metaphysically distinct personal level or even one of much epistemological import). Thinking about the overarching nature of cognition, as a generic kind, has also catalyzed a model-based theory of natural kinds, the „tweak-and-extend" theory, found in some of my most recent work.

How do you perceive the relationship between philosophy and science? And what do you think is the value of interdisciplinarity?

I think the relationship has many facets. Philosophers familiar with scientific results and theories can interpret that work for other philosophers and for the educated public, by bringing it to bear on issues of long-standing philosophical and popular interest. Some such work can be quite revolutionary, changing our everyday conception of selves or of social interaction, for example. At the same time philosophers can play the role of theorists in the sciences themselves, offering precise accounts of foundational concepts (such as that of representation) or attempting to help straighten out inconsistencies in the use of those concepts. To my mind, though, what's most interesting, challenging, and rewarding results from a kind of freedom philosophers have to survey the landscape and attempt to provide an integrated and coherent vision of some large domain, identifying a pattern that appears only when one considers a wide variety of work being done on different sub-topics in a field or in different scientific fields altogether.

What sorts of gadgets do you usually use in order to "extend" your own mind? We mean both science-related extensions as well as the ones that you use solely for entertainment. Do you read a lot of printed books? Do you still write using pen and paper, and if so, then how often?

I do most of my writing on my laptop, and I use my iPhone quite a bit. I use pen and paper almost every day, mostly for doing quick calculations, making lists, or writing personal reflections and notes. I read only paper books, although I do often read journal articles on the computer screen. I'm even more of a Luddite when it comes to entertainment. In that domain, I don't use any "mind-extending" technologies (although it depends on how liberal an approach we take—is a guitar mind-extending technology?).

What educational and academic choices guided you to the subjects you are currently interested in? Did you experience any dramatic change in research interests during your career? What did you imagine your future would turn out to be like when you were a teenager?

I was very much interested in psychology and the human condition when I was young, but largely through literature, history, movies, and songs—in other words, from a more humanistic perspective. But as I began taking college courses, analytic philosophy seemed to promise the most enduring or abstract truths, and it also drew on my longstanding fascination with analytical puzzles and games. I think those inclinations ultimately led me to a scientific approach to the exploration of content, consciousness, and cognition. The sort of philosophy that offered the deepest and less purely speculative insight into the mind was oriented toward the sciences. I might add that in my first philosophy of mind course, the professor (Charles Marks) assigned Nisbett and Wilson's 1977 paper, "Telling More Than We Can Know: Verbal Reports on Mental Processes," which instilled in me a healthy skepticism regarding our "pure" philosophical intuitions about the workings of the mind.

My current research interests grew initially out of my fascination with logic and philosophy of language. These latter interests turned me toward the mind at the very end of my undergraduate education when I encountered externalist semantics for mental states. Then, in the summer after I finished my bachelor's degree, I attended a seminar on mental content co-taught by Scott Soames—who was at that time a serial visitor in my department—and Charles Marks, my department's resident philosopher of mind. I had been thinking extensively about the indeterminacy of reference and the causal theory of reference, and the way some of these issues connected up skeptical views in epistemology and philosophy of science. Then, in the seminar with Soames and Marks, I saw how to cast many of these issues as questions about mental content and the mind's causal connection with the word. It was in this course, too, that I had my first substantive exposure to Fodor's work and began to think of philosophy of mind as deeply empirically informed.

When I was a teen-ager, I wanted to be a musician, a folk rocker like Bob Dylan or Neil Young—and I wanted play lead guitar like Jerry Garcia. But I was realistic with regard to my career options. I had intellectual leanings and abilities and assumed they would lead to stimulating remunerated work, probably to do with politics or history, that would in some way put me into contact with great texts of history and the community of human thinkers. But, to be honest, none of these thoughts took the form of concrete commitments or plans when I was a teenager. I was carefree and idealistic. I wanted to play music and read books and think deeps thoughts, and my attitude was, well, whatever happens happens. (I must admit that I encouraged my own children to give their adult futures more careful consideration!)

Bibliography

Anderson, M. L. 2010. Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, 33(04): 245-266.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20(04): 723-742.

Clark, A. 2013. Gesture as Thought? Z. Radman, ed. *The Hand, an Organ of the Mind: What the Manual Tells the Mental*. MIT Press.

Fodor, J. 2011. An Interview with Jerry Fodor. *Rutgers Philosophy Newsletter* (Spring): 8-10 http://philosophy.rutgers.edu/dmdocuments/Rutgers%20philosophy%20newsletter%20(Spring%202011).pdf (10.11.2012)

Lakoff, G., & Johnson, M. 1999. *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books.

McClamrock, R. 1995. *Existential cognition: Computational minds in the world*. University of Chicago Press.

Nisbett, R. E., & Wilson, T. D. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3): 231.

Rupert, R.D. 2009. *Cognitive Systems and the Extended Mind*. Oxford University Press.

Shapiro, L. A. 2011. *Embodied cognition*. London: Routledge.

Williamson, T. 2013. How Deep is the Distinction between A Priori and A Posteriori Knowledge? Casullo, A., & Thurow, J. C. *The A Priori in Philosophy*. Oxford University Press.

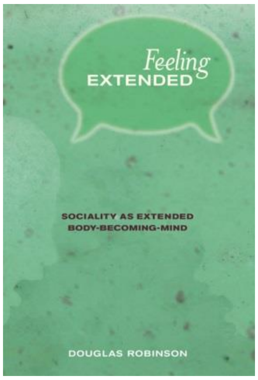
book reviews

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0012

Feeling Extended

A book review of *Feeling Extended: Sociality as Extended Body-Becoming-Mind*

Author: Douglas Robinson
Publisher: MIT Press
Release Date: 2013
Number of Pages: 256



Patricia Grosse
Department of Philosophy,
Villanova University
patricia.grosse()gmail.com
Received and accepted September 2014; published Autumn 2014.

In their (in)famous work, "The Extended Mind" (1998), Andy Clark and David Chalmers introduce the Extended Mind Hypothesis (henceforth, EMT) by means of a famous example: Otto and his notebook[61]. This notebook ought to be considered as an extension of Otto's mind, they claim, as he uses it as a functional equivalent of certain cognitive capacities that are normally lodged inside the head. Clark and Chalmers have to do the hard work of laying the groundwork for getting cognition out of the head. However, EMT is missing something that Clark and Chalmers find pernicious. In fact, if this review were a proper space for joking, I would say that Clark and Chalmers do not want to include into their theory what many men do not like to even talk about: their feelings. That is, they seek to avoid the "qualia trap," by which they mean the question of how we feel about our experiences (qualia). In his recent book *Feeling Extended: Sociality as Extended Body-Becoming-Mind*, Douglas Robinson introduces "feeling" back into the extended mind in a compelling and effective way. Currently Chair Professor of English and Dean of the Arts Faculty at Hong Kong Baptist University, Robinson's varied research interests and publications are all found by the common thread of inter- (and perhaps intra-) communication. Robinson is in the difficult position of strad-

[61]I would like to thank Georg Theiner for his helpful comments and suggestions on an earlier version of this review.

dling two worlds: Materialism and Idealism. He writes of himself in the Introduction to *Feeling Extended…*, that he is "a materialist who recognizes that everything we know about material reality is a quale, and an idealist who recognizes that qualia are human groups' ultimately inadequate attempts to represent and control material reality" (15).

As a general note, the author's style is florid and convivial. This book is truly a pleasure to read (or at least feels that way). The author's personal accounts are a pleasant surprise—his novel method of counting laps while swimming, a fight with his (now ex-) wife in Note 12 of Chapter One that leads to a broken fender and a psychically wounded arm, his feeling at home with the Spanish language despite his lack of ability to communicate in it. He is able to paint a picture so vivid that the reader too must find herself victim to the conative force of his writing. Included in this text is a delightful Appendix called "Liar-Paradox Monism." This Appendix should be read as soon as possible upon picking up the text. In it Robinson addresses Chalmers' Hard Problem of Consciousness ("the problem not only of *how* but of *why* certain physical structures give rise to consciousness (or experience, or qualia)" (Robinson 179)). That is, we cannot get at an understanding of qualia from the standpoint of the "real," physical world. The two options we are left with to make sense of qualia, on Robinson's account, are "a naturalistic or interactive dualism that posits two realms, the physical and the mental, and builds bridges between them, and a panpsychic monism that explains the physical as proto-experiential" (179). Robinson's answer to the Problem is to lie—a philosophical position that he has playfully termed "Liar-Paradox monism (LP monism)." By giving an analysis of Oscar Wilde's "The Decay of Lying," Robinson offers a way out of the problem by postulating the possibility of maintaining a position while acknowledging that one may be lying to oneself or others about the veracity of said position. That is, "what LP monism does … is embrace the dissonances, embrace the complex phenomenally and rhetoricity of our engagement with the world, and so offer a truer account of the world as we experience it" (203). If one were to make sense of this Appendix one would make sense of the overall force of the text. We cannot give a "real" account of the world—we are always-already in a situation that is determined by myriad factors beyond our capacity to fully understand, though it is possible that we might tell ourselves some sort of sense-making narrative based upon our experiences. But, "I could be lying to myself" (206). Robinson gives a clearer example of what he means by the liar-paradox in Chapter One: "my computer, who is my friend, sometimes unaccountably acts like an inanimate object. As a result, I form a belief that it probably *is* an inanimate object, and keep provisionally contrasting that belief with my strong *feeling* that it is alive and friendly and there to help me do my work till I am willing to say, tentatively, that probably only *seems* to be alive" (66). That is, my interaction with my computer is as-if it were really alive and not merely a mechanical extension of myself. While I know "intellectually" that I am the one who set up applications on my computer that make writing easier and also that I am the one who interacts with the interface meaningfully, when my computer breaks down I am forced to see the lie I've been telling myself.

In Chapter One of his book, Robinson gives an account of the central concepts of the EMT. He undergirds his argument with Hegel's philosophy of mind regarding the interaction between mind and tools. While the bulk of the work done by proponent of within Philosophy of Mind has been to articulate and defend the claim that our consciousness—at least, parts of our cognitive apparatus—extends into the world, outside of the brain and even outside of the skin, Robinson offers an "as-if" (to borrow from Damasio 1999) EMT—we feel our consciousness as-if it were extended. Though the chapter is titled "Inside Out," Robinson suggests that a fuller title would be "Inside Out and Outside In" (34). In elucidating these two points I will turn to the examples Robinson uses: his two wives. In "Outside in," which refers to one's feelings towards the things outside of us, he gives the example of his own experience of learning Russian: Russian was a foreign tongue to him until he heard his second wife "speaking loving Russian" to their daughter (53). Once he heard the language in this loving context, he *felt* more at home with it. He writes, "If learning is largely an outside-in operation—the internalization of words, numbers, images, ideas to which we are exposed outside our heads—there must be a process by which things that are alien to us because they are outside us gradually become part of our affective-becoming-conative 'tissue'" (55). By "inside out," Robinson engages with the Hegelian notion of tools and cognition as extended desire (or, "interactive affect-becoming-conation" (55)). There is a revealing section in this chapter in which he discusses "Proprioceptively extended cognition," whereby the individual throws out their mind-map, for want of a better expression, onto the world at large. Robinson and his first wife, evidently, had a contentious relationship. One argument ended with their car fender being dented. Although he was outside the car watching the accident unfold, he "felt a stabbing pain in my left shoulder. My nervous system, mapping my body onto the car's, lying prone with my wife "steering" me, simulated in my body the "pain" "felt" by the car's body. Antonio Damasio (1999: 80) calls the neuronal system that makes this simulation possible the "as-if body loop," but mainly means by it our bodies' tendency to mimic the body states of other *people*. The possibility that we also simulate the "body states" of inanimate objects would be a materialist explanation of the pathetic fallacy" (Robinson 2013, n.1.12, 211). The extension of one's body to an inanimate object is intriguing in its obviousness. When one sees one's car with scrapes on the side after a particularly terrible parking job, one winces, regardless of whether the scrape happened yesterday or three years ago. Why is that?

In Chapter Two, "Language as Cognitive Labels" Robinson engages with Adams and Aizawa, who so dearly seek to keep cognition in the head. He offers "a counter model (to Fodor's language-of-thought hypothesis and the Computational Theory of Mind) by tracing the actual emergence of thought out of embodied (affective-becoming-cognitive) communication with others" (28). Robinson seeks to problematize what he sees as the "binarization in terms of cognition and noncognition" of the "spectrum of thought and language" by appealing to "fuzzy logic" (78). That is, Robinson claims that Adams and Aizawa do an injustice to the inexorably intertwined nature of thought and language. To put it plainly, Robinson sees a difficulty with the claim that thought stays in the head and affects language and is not in turn affected by our own use of language or other transcranial forces (i.e. the language of other people). Following Robinson's example in this chapter, I say something rude to my friend. He asks me if I am hungry and therefore "hangry" and I say "Yes." He offers a cookie and things are resolved. Where did the realization occur in me - the realization that I was angry because I was hungry? It was only upon further reflection *due to his comment* that I realized that I was indeed being aggressive because my blood sugar was low and I desperately needed to eat something. The force of this argument in this chapter is thus: it is silly to divide up this interaction into "discrete chunks labeled 'cognition' and 'noncognition'" because that would be like trying to separate the "white" from the "rice" (84). Basically, Robinson claim is that there needs to be more acceptance of the fact that we not only are embodied creatures who think, but also thinking embodied creatures who interact with and are affected by the world and others.

In Chapter Three, "Language as Conative Force," Robinson critiques the limited notion of language as merely verbal labels (what Robinson considers the rationalist philosophy of language) he outlines in Chapter Two. Robinson delves into what he calls "the focal claim" of his book: "that *something* connects us, non semantically (non propositionally)—that sociality really is a form of extended body-becoming-mind" (28). This connection can be seen in the simple fact that there is much more to communication than simple words. The meaning of a simple sentence is closely bound with the performance of that utterance. Anyone that has inadvertently started a fight with a loved one through text message knows the importance of tone and body language all to well. And how are we to describe this connection between speakers that makes them able to understand each other's meanings? Throughout this text Robinson brings key figures in the history of philosophy, religion, and literature into conversation with more "scientific" theories of cognition and communication in order to better develop his project. In this chapter, he develops his theory of conative *force* by engaging with Jacques Derrida's *iterability*, Mikhail Bakhtin's *internal dialogism*, and Pierre Bourdieu's *habitus*. In doing so, Robinson is better able to engage with "the possibility that there are other channels of communication, especially some sort of affective-becoming-conative *force* that is often directly verbalized as speech acts, but, as indirect speech acts suggest, need not be verbalized in order to be transmitted" (118).

These forces which are passed person to person are qualia, Robinson claims in Chapter Four, "Qualia as Interpretants." When I yawn I cause my friend to yawn; this transsomatic, transcranial activity is an interpretant that has a "qualia" force. As we all know, yawning is highly contagious. I am yawning as I write this sentence—not because I am tired, but because I am contemplating just the word itself. When I gave the example of yawning in order to describe the nature of the interrelation between qualia and interpretants with a colleague over the phone, she yawned as well. Transcraniality thus need not be local. Not surprisingly, Robinson relies on the foundational work of Charles Sanders Peirce to gird his views. What Robinson has to contend with is the seeming absurdity of qualia. Is this because men don't like to talk about feelings? Or at least take them as seriously as they do more rational things? Daniel Dennett certainly doesn't, when he calls us to "'quine' (deny the existence of) qualia" (Robinson 119; see Dennett 1988). However, by denying qualia as mere phantasms, Dennett, Clark, et al. ignore the force that these qualia have on bodies-becoming-minds. When arguing with a lover, after all, it is best to use "subjective" language that swerves around "objective" reality. As a case in point, "I felt like you don't respect my work when you ask me to watch our child when I'm trying to write," is a more effective form of communication than "you don't respect my work because you asked me to take care of the kid." In Robinson's words, "The notion that mind-as-qualia is somehow too ethereal to guide or steer the body is sheer atavistic Cartesianism" (145). A theory of cognition that doesn't account for affect, or the fact that human beings are bodied and always surrounded by other bodied human beings, is unavoidably negligent and lacking. How I experience the world ought not to be discarded, and to do so would be to claim a strong distinction between the purity of mind and the muddiness of feeling, emotional bodies.

The affective embodiment theory Robinson promotes need not be completely baseless and outside of contemporary research in neuroscience. In Chapter Five, "Empathy, Face, and Ritual," Robinson engages with social neuroscience in order to further explicate his thesis. "What this empirical research strongly suggests," says Robinson, "is that we don't necessarily *know* what others are feeling, but we do tend to *feel* what they are feeling" (153). Once Robinson lays out the "empirical evidence" for bodily and mental representations as qualia, he turns to speculation. In an extensive section on the Ancient Greek—by way of Aristotle—notion of *doxa*, Robinson notes that "face" is a useful alternative translation for *doxa*, which my Ancient Greek lexicon enumerates as the opinion which others have of one, estimation, reputation, credit, honor, glory, etc. *Doxa* is related to one's situated identity. That is, the individual is interested in maintaining face, maintaining her ability to have a certain social standing, by any means necessary. And this "face" she wishes to keep is dependent on how others perceive her. This account of Ancient Greek *doxa* is a lovely way to demonstrate his general theory of bodies-becoming-mind converting qualia into collective behavioral pressure (145). Let us unpack this: qualia is how we feel about things in the world. Some call this fleeting, ethereal, not "really real" But how we feel about the world informs how we think and act at a deep level, perhaps even all the way down. We are bodies-becoming-mind, in Robinson's view, because of the complex interweaving and enmeshing of cognitive and non-cognitive. Indeed things are so mixed up between body and mind that separating the two in order to contemplate and catalog either ends up in poor science. Robinson thinks that this enmeshing is also extended to societal interactions - as already in a social group, we convert qualia - how we feel about the world - into collective behavioral pressure. An example of this can be seen in Robinson's discussion of the conflation of feeling and looking in the context of mirror neurons: "what enables us to distinguish my moment from yours, the individual from the group, is in fact a secondary cognitive (meta)quale that is belatedly imposed on transcranial proprioception." (166). To Robinson, feeling caught up in an event such as the opening ceremony of the Beijing Olympics is not to be unconscious and unmindful, it is a distinctive mode of being—body-becoming-mind. Thus the "I" in the *Cogito* is an I that is after, an I that is already extended body-becoming-mind, an I that thinks it has stripped itself bare of everything but the essential. But to be human is not to be a mature mind fleshed in a body sitting comfortably by a fire, but to a being who once was an infant, who learned to talk, and, hopefully, has loved and been loved in return. To be human is to be already enmeshed in the world.

Perhaps the most shocking claim of the book (even more shocking than Chapter Four's nearly *ad hominem* attacks against Andy Clark, which include accusations of Clark inhabiting Plato's cave) is a very innocent remark in Chapter One concerning our use of tools: while extending mind to tools makes Otto's day easier and his capacity for cognition greater, not all tool use ends up with sharper users. It's the dark secret of the EMT, on Robinson's view, that the very nature of the human-tool interaction can lead to "brain-fried zombies" (42). That is, not all extension of mind need make one more intelligent, it may make one duller. This is not at all to undermine the EMT, as while Google may make out memory retention duller, it can arguably lead to more successful being-in-the-world (see also Carr 2011 and Sparrow & Wegner 2011).

Robinson seeks to inhabit a between place between pure extension and pure "flesh-box." In doing so he does not do himself any favors with potential readers thoroughly entrenched in this or that camp within the Philosophy of Mind. Perhaps his insistence on not denying or fully ascribing to the EMT is not convincing enough for either hardline externalist (e.g. Clark) or internalist (e.g. A&A) camps. However, doesn't his between position relieve us of a great burden: the burden of proof? If we know at base that cognition *feels* extended, is that not enough? Robert Rupert, a critic of the EMT, argues in his *Cognitive Systems and the Extended Mind* that the thinkers such as Wheeler and other neo-Heidegerrians are mistaken in their view that we should take into account our alreadiness, the fact that we are always-already in a culture, when positing accounts of the extension of mind. He claims that this phenomenological view "suffers from an obvious and crippling foundational problem that the filling in of further details does not alleviate. The Heideggerian view offers no explanation of how the human internalizes social norms or of how social norms come into existence" (Rupert 2009: 162) Robinson gets around the difficulty of coming up with a proper syllogism to withstand Rupert's critique by rejecting the need for strong, hard science to back him up. There are plenty of studies he can point to in support of his claims (which he does at length in Chapter Five), but his appeal to fuzzy logic as well as LP monism relieves him of that burden of proof.

In *Feeling Extended…*, Robinson misses the opportunity to bring Feminist Epistemology in conversation with EMT and Affect Theory. In Chapter Four, Robinson really drives the stake into the heart of the primacy of science: "Empirical testing, after all, might reveal the extent to which empirical science rests on the gossamer foundation of qualia" (Robinson 143). He brings Thomas Kuhn's work into the conversation as a way of problematizing the way in which "Science" is enshrined in our culture. This distrust of the "objectivity" which empirical science lays claim to has long been discussed in Feminist Epistemology. The moves Robinson makes, while interesting and important, are not earth-shattering when read against such figures as Donna Haraway, Sandra Harding, or Susan Bordo, to name but a few. That being said, it is not possible to engage with the entire history of philosophy in a given text, even for a text that engages so heavily with the literature spanning several fields, disciplines, and theoretical approaches. Robinson has cleared the ground—if you could forgive the Heideggerian phrase—for research that remains to be done in this area, i.e. the linking and critical interrogation of theories of Extended Mind from the perspective of Feminist Epistemology. After all, what is missing in classical approaches to Extended Mind? How we feel, how we interact with others, basically, how "we" are in the world.

Feeling Extended… is well-researched, well-written, and one of the more enjoyable texts one may come across in its field. It is - on the face of it - an answer to the 'qualia trap' that Clark and Chalmers so eagerly aim to avoid, but much more sophisticated than that. Douglas Robinson opens the way for new work in the field of 4E+A by calling a spade a spade, a feeling a feeling.

References

Carr, N. 2011. *The Shallows: What the Internet is Doing to Our Brain*. W.W. Norton & Company.

Clark, Andy, and David J. Chalmers. 1998. The Extended Mind. *Analysis* 58: 7-19.

Damasio, Antonio R. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* San Diego: Harcourt.

Dennett, Daniel C. 1988. Quining Qualia. In *Consciousness in Modern Science*, ed. A. Marcel and E. Bisiach. Oxford, UK: Oxford University Press.

Robinson, Douglas. 2013. *Feeling Extended: Sociality as Extended Body-Becoming-Mind*. Cambridge, MA: MIT Press.

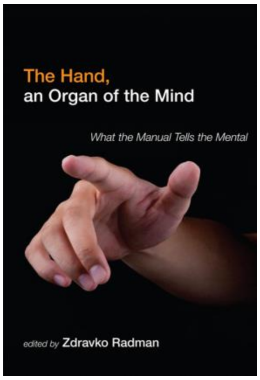
Rupert, Robert. 2009. The Experience of Extension and the Extension of Experience. In *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.

Sparrow, B. Liu, J. & Wegner, D.M. 2011. Google Effects on Memory: Consequences of Having Information at Our Fingertips. *Science*, 333: 776-778.

AVANT, Vol. V, No. 2/2014
ISSN: 2082-6710 avant.edu.pl/en
DOI: 10.26913/50202014.0109.0013

The Hand and Cognition… and Intersubjectivity, Agency, Culture, and More

A review of *The Hand, an Organ of the Mind: What the Manual Tells the Mental*



Editor: Zdravko Radman
Publisher: MIT Press
Release Date: 2013
Number of Pages: 464 pages

Christopher Drain
Department of Philosophy,
Villanova University
cdrain1()villanova.edu
Received and accepted September 2014; published Autumn 2014.

The Hand, an Organ of the Mind, edited by Zdravko Radman, consists of seventeen essays by an impressive array of philosophers, psychologists, and cognitive scientists, as well as a foreword by Jesse Prinz and a postscript by haptic artist Rosalyn Driscoll.[62] As the title suggests, this collection concerns the role of the hand in cognition and consciousness, and aims to occupy a novel middle space in the polarized world of philosophy of mind and consciousness studies. As Radman points out, there has since Descartes been a methodological dualism in studying the mind, with variations of the mind-body binarism taking form in the contrasts of "organic versus inorganic, mind versus matter, 'I' (or self) versus 'it' (brain), cognitive versus motor, internal versus external, subjective versus objective"… the list goes on (Radman: xix). The place and function of the hand, however, does not fit well on either side of such a schema, and it is because of this fact that handedness has been neglected and under-researched in theories of mind and cognition. Indeed, as the authors of this collection demonstrate time after time, the hand serves as a unique theoretical vehicle that, if not entirely bypassing such traditional dualisms, at least complicates them insofar as it often shows itself to play a mediating role in the

[62] I would like to thank Georg Theiner for his helpful comments on an earlier draft of this review.

relationality between intra- and extra-cranial cognition, agency, and mindedness. Just some of the most stark claims in this work include: that not only is the hand a key ingredient in communicative practice in general (in gesture and writing), but that the specific gesture of pointing is integral to the phylogenetic and ontogenetic development of symbolic thinking all together (Cappuccio and Shepherd); that the hand and its tactile receptivity provide the basis for inter-empathetic subjectivity (DePraz); that the hand serves as a facilitating device for the cognitive integration of cultural artifacts via the redeployment of less specialized Pleistocene neural functionalities (Menary). This volume makes a hard case for the fact that the hand, as an organ through which we not only explore our social and objective world but also produce it, is integral to the human experience.

The essays in this text are grounded in several traditions and disciplines, ranging from neurophysiology and developmental psychology to contemporary robotics and artificial intelligence, and from the history of philosophy and the phenomenology of the classical German and 20th century French varieties to cognitive science and the "4E" informed branches of philosophical psychology.[63] For a reader who wants a state of the art, cross-disciplinary account of the hand, then this variety is a virtue. But a researcher with a specific problem to tackle may not benefit from such a wide theoretical scope. This latter point, however, should not be taken as any substantial criticism. If there is minor room for complaint here, it is regarding the text's organization, in that the essays sometimes seem unevenly distributed. The book is organized thematically, with each section containing works from any number of the aforementioned theoretical holds. However, some sections are longer than others, the first and second clocking in at around one-hundred pages each, while the third is under fifty and the sixth and final section (containing only a single essay and a postscript) is under thirty pages. Moreover, some sections seem to be grounded less pluralistically than others, and at times, it seems arbitrary why one particular piece shows up in one section rather than another. For instance, the first, longest section, "Hand-Centeredness" contains essays based more in laboratory work than philosophical reflection and argumentation. That's fine. There is nothing wrong with the content of those works. But to the non-specialist, the section is perhaps the most tedious—it certainly front-loads the text with a body of work that while quite rigorous and exact, is less reflective than the rest of the book. There were several essays based squarely in the phenomenological tradition that did not warrant their own section, though they could have had the editor decided as much. This is not a bad thing, how-

[63] The "4E" label includes the programs of extended, embedded, enacted, and embodied cognition that find themselves united insofar as they eschew the intellectualist paradigm that treats mentality as inherently representational and/or computationally representable. See Menary 2010.

ever. They were placed at appropriate junctures thematically linking them to works of a different tradition, providing measured counterbalances. And while the essays of section one are all "Hand-Centered," the same can be said of basically every essay in the collection. What does link the particulars of the first section is the disciplinary affiliation and background of its authors within experimentally based psychology and neuroscience. But I see no prima facie reason why certain contents of section one could not have been dispersed among other sections. The conclusions of those works were certainly philosophically relevant. As such, they could have been appropriately placed alongside more straightforward philosophical reflections, the claims of which being either buttressed or challenged by the results of the experimental psychologists. This would have the positive result of highlighting the philosophical implications of the chapters of section one, implications which as it stands are not as ready-to-hand as they could be in a different arrangement. This is a minor quibble, however, as I suspect not many people will read the text straight through. Given the interdisciplinary scope of the work, many will pick and choose according to their respective interests. But blocking off essays from one particular field might tempt a hitherto disinterested reader to continue along in her or her ignorance.

As mentioned, this is a lengthy work. There are seventeen essays divided among six sections, totaling over four-hundred pages. Unfortunately a full review of each essay would require more space than allotted here. But in the following I will summarize each work, highlighting its central themes. When called for, a more extended discussion will be presented.

1. "Hand-Centeredness"

While most of the essays in this volume are hand-centered, this section focuses specifically on the neuropsychology of hand-centered activity. Jonathon Cole's essay opens the discussion with a survey of relevant case studies of subjects with either partial or total sensory loss. Following the rare cases of those afflicted with neuropathy syndrome (where there is an "acute loss of proprioception and in most cases, touch"), Cole describes subjects' whose loss of proprioception results in cases of limb kinetic ataxia, or the loss of any controlled movement, even though the motor nerves themselves function normally (Cole, 7). Some subjects were able to relearn movement, however, but only with painstaking concentration to details that normally go unnoticed by unafflicted subjects. Such relearning, however, is only possible by constructing elaborate plans through trial and error. At the very least, Cole shows that the intellectualist paradigm of agency as the product of a ratiocinating and "central executive" planner only seems to obtain in rare and aberrant cases and that normally mundane actions are the result of an embodied coping mechanisms. While the essay at time concerns bodily movement in general, Cole does focus on cases of relearning gesture, which seems comes easier than other types of bodily movements. Citing David McNeill (2005), Cole concludes that apraxic subjects experience less hardship in these cases because the system that controls the hands during gesture action is linked with a "thought-language-hand" system that differs from the system controlling instrumental action (Cole: 17).

Andrew Bremner and Dorothy Cowie focus on the ontogeny of *representations* of the hand, theories of which, they point out, are relatively obscure in developmental psychology when compared to more mainstream accounts of the role of hand action in cognitive development. They begin their essay by outlining the unresolved debate concerning whether hands (and environmental interaction in general) or inborn intelligence have priority in cognitive development. Piaget (1952, 1954) serves as an early representative for the hands-priority thesis, while the intellectualist position is finds support in Spelke et al. (1992) with a "core knowledge" approach that effectively argues that knowledge of how to interact with the environment is the result of a phylogenetic inheritance. Both theses turn out to be limited, however. Piaget undercuts the role of the hands and body, viewing them as ultimately hindering abstract perceptions and action schemas, whereas the "core knowledge" approach never calls into question just how such knowledge is able to be enacted. The authors turn to current research on infant hand representations, tarrying between top-down and bottom-up theories of body-schematization and internal body modeling. They conclude with a compromise: while infants do seem to start out with core knowledge, it is still the case that body schemas develop through manual interaction with the environment up until a much later age.

The last two essays in this section are respectively devoted to hand-centered space and peripersonal awareness. Nicholas Holmes' essay "Hand Centered Space and the Control of Movement" makes the case for a hand-centered visual representational field. To demonstrate that there are neuronal representations of hand-centered space, Holmes outlines several experimental situations, from the rubber-hand illusion to cross modal extinction and congruency tasks. Holmes shows that the hand occupies a unique role in in "determining a participant's ability to detect, discriminate, or pay attention to visual or somatosensory stimuli" (Holmes: 60). But Holmes points out that most of such experiments only study the hand in passive situations, not taking into account the primary function of hand-centered mechanisms—that of action-centered attention. Holmes closes his essay arguing that the most basic functions of hand-centered representations are the (evolutionarily relevant) actions of hand defense movements and desire-directed movements toward target objects. Continuing with the notion that the hands occupy a primary role in visual-representational action orientation, Matteo Baccarini and Angelo Maravita argue in "Beyond the Boundaries of the Hand: Plasticity of Body-Space Interactions following Tool Use" that during tool use "we can modify our relationship with external space in terms of body/space representation" (Baccarini & Maravita: 82). Their main claim is that one's body schema is altered during intentional and effective tool use, which itself changes the potentialities for action in one's peripersonal space.

2. "Togetherness in Touch"

The second consists of literature focused on the intersubjective relevance of touch and of more ontologically themed essays on touch and perception and reality (the section's title is apt in this second instance if we think of "togetherness" asocially as an "ontological binding"). It commences with Harry Farmer and Manos Tsakiris' "Touching Hands: A Neurocognitive Review of Intersubjective Touch." This essay is an informative take on recent research into intersubjective touch, and includes sections surveying findings from evolutionary, psychological, and neuroscientific perspectives. The authors move from the social relevance of primate grooming patterns to contemporary socio-psychological factors that modulate human intersubjective touch, from gender and age to setting and type of touch. They then turn to the importance of touch in infancy for the development of social and empathetic intelligence. The essay closes with a review of the neural bases of intersubjective touch.

The two middle pieces of this section are related insofar as they deal with more general ontological and perceptual implications of touch. Matthew Ratcliffe's "Touch and the Sense of Reality" turns to classical phenomenology to question the hegemony of vision in the philosophy of perception, while Filip Mattens' "Perception and Representation: Mind the Hand" explores the subjective reasons that touch succumbs to visual hegemony in the first place. Ratcliffe claims that even when touch gets its due credit, it is often only the hands that get mentioned, and the more general form of "background touch" is overlooked. Such a sense, it is argued, is primary in connecting us with the environmental significance constitutive of our "world" (or *Weltheit* even though Ratcliffe doesn't specifically use this term). Discussions that focus only on the exploratory and actionable aspects of touch miss the sense of world-disclosure that undifferentiated touch, as underlying the perceiver-perceived relationship, fundamentally grounds. It is this touch that can best "illuminate the sense of commonality (…) that is presupposed by the possibility of encountering anything as 'there'" (Ratcliffe: 139).

Filip Mattens' piece is an appropriate accompaniment to Ratcliffe's. Mattens discusses the ways in which the hands deceive the intellect into upholding the "epistemological credo" that touch is primarily, like eye-sight, an object-sense (Mattens: 159). Mattens argues that this tendency obscures the fact that touch is primarily not a perceptual sense but a "vigilant sense"—"the tactile sense serves an organism not for touching but for sensing that it is being touched"

(Mattens: 166). The hand however reverses this logic insofar as its exploratory movements are often linked with what is seen; touch becomes responsible for retrieving objective information about the world. At this point the reader can see the usefulness of this collection—already there is opportunity for an encounter between Mattens work and Holmes' work on action-centered attention. Also, one is curious to see what Ratcliffe would have to say about Matten's talk of touch as primarily a mechanical encounter with an object's shape—a direction of theorization that he criticizes as ignoring the intersubjective significance of the world as disclosed through touch.

Moving on, Natalie Depraz' essay "the Phenomenology of the Hand" picks back up on the notion of intersubjectivity and the relevance of touch for inter-empathetic attunement. Her essay should be appreciated not only for highlighting this often under-theorized intersubjective aspect of touch, but also for providing the volume's most thorough philosophical history of hands and touch. Classical phenomenology aside, there is sometimes a noticeable paucity of historical awareness in this text and it would have been nice to see the authors cite and reflect on their philosophical forebears more often (e.g., the editor's claim in his own essay that "we are *both subjects and objects of our own doing*" seems straight out of Marx, but Marx is never mentioned) (Radman: 386).[64] Depraz however is acutely aware of history of reflection on the hand. She constructs a narrative of the dialectic between the theses "that humans are intelligent because they have hands" and "that humans have hands (i.e., can use their hands effectively) because they are intelligent." These positions seem to first arise with Anaxagoras and Aristotle but Depraz shows they respectively resurface in Engels and Bergson (such an opposition indeed motivates Bremner and Cowie's article mentioned earlier). What Depraz goes to show is that in the history of philosophy the hands are only a matter of interest as they relate to knowledge and mastery—regardless of whether the hands are the source of intelligence or vice versa, they maintain the function of facilitating knowledge of the world or of practically altering the world to our benefit. The function of the hand as facilitating intersubjective recognition is passed over in this respect. Depraz spends the remainder of her essay surveying phenomenological accounts of touch and inter-empathetic awareness in the works of Husserl, Merleau-Ponty, Levinas, Sartre, Danis Bois, and Michel Henry.

[64] Compare Marx: "When real, corporeal (*leibliche*) *man* (…) establishes his real, objective *essential powers* as alien objects by externalization (*Entäußerung als fremde*), it is not the establishing (*Setzen*) which is subject; it is the subjectivity of the *objective* essential powers whose action must therefore be an *objective* one. An objective being acts objectively, and it would not act objectively if objectivity were not part of its essential nature. It creates and establishes only objects because it is established by objects, because it is fundamentally *nature* (*weil es von Haus aus Natur ist).* In the act of establishing it therefore does not descend into 'pure activity' to the *creation* of *objects*; on the contrary its *objective* product simply confirms its objective activity, its activity as the activity of an objective, natural being" (Marx 1992: 389).

3. "Manual Enaction"

Most of the latter half of this volume is devoted to essays working within the various paradigms of "4E" cognition. Prominent theorist of embodied cognition Shaun Gallagher starts off this section. His "The Enactive Hand" covers three areas of concern for the hand in cognition, all in line with his thesis that rationality is proximally and for the most part action-oriented. The first section regards the complicated relationship between the hand and vision. Gallagher demonstrates through laboratory evidence (notably the rubber hand illusion) that although vision seems to trump the hand in certain settings (we are fooled by sight alone into thinking that the rubber hand is ours, evenly experiencing experience tactile sensation that is objectively not ours, even though it phenomenally seems so), we are not so fooled when motor-systemic awareness comes into play. Citing Iverson and Theleon (1999), Gallagher goes on to argue that the hand, vision, and neural circuitry take part in a holistic operation. Hand movement is neither a bottom-up environmental emergence of rational action nor is it a top-down determination of movement. Rather, the hand and brain exist in a "single integrated cognitive system" implicative of a "reciprocal unity of feedforward-feedback processes in which the hand and the brain form a dynamic system that reaches into the world" (Gallagher: 213). Gallagher goes on in the next sections to describe hands in their practical function of worldly engagement, which reads well along Baccarini and Maravita's piece on tool-use and body-space interactions. Gallagher closes by reviewing the social and communicative relevance of hand movements.

"Radically Enactive Cognition in Our Grasp," Daniel D. Hutto's contribution to the collection, makes a case for rejecting the representationalist understanding of the mind, or the idea that all thinking consists in rational deliberation, where thoughts are structured mental representations that have propositionally representable truth values which are computationally manipulatable. Hutto does not claim that such ratiocinative thinking does not ever occur—in complex future planning it obviously obtains. His point, rather, is that deliberative cognition only occupies a small domain of intelligent mental functioning and representationalist theories ignore a more primary form of cognitive activity in which "the specified body and environmental factors are fully *equal partners* in constituting the embodied, enactive intelligence and cognition of (...) artificial and natural agents" (Hutto: 231). Hutto looks to the hand to make his case, since it seems that much of successful manual activity does not rely on rule following and propositional attitudes but rather on spontaneous alterations that are situationally determined. If intellectualist assumptions were to obtain, then it would be the case that manual activity, and indeed all bodily movement, would result from the type of calculated planning the Jonathan Cole describes in patients with limb kinetic ataxia—which seems to be contradicted by psychological findings. Hutto goes on to provide an intricate discussion of radical and conservative takes on extended cognition, the latter not fully jettisoning representations insofar as its supporters maintain that there exist mental representations encoded in bodily formats, or "action oriented representations." A crucial issue arises concerning the nature of informational content, which presents a stumbling block for this more conservative approach. After a novel series of argumentations, Hutto concludes that a cognitive system is such that its sensory systems do convey information, but not in any sense of passing on *meaningful* or *contentful* messages that articulate inner mental representations. Rather, "cognition activity involves complex series of systematic—but not contentfully mediated—interactions between well-tuned mechanisms" (Hutto: 248).

4. The Gist of Gestures

As the title suggests, this section deals with the cognitive role of gestures. Andy Clark's "Gesture as Thought?" makes the case that rather than being used to communicate already formed thoughts, gesture primarily plays an active causal role in thinking. Clark, keeping with his role as co-originator of the "Extended Mind Hypothesis," (EMH) argues that gesture is part of a coupled intra/extra cranial system, and constitutes an organismically extended mode of thought. Building on McNeill (2005) and Gallagher (2005), Clark describes how gesture may function in a self-stimulated feedback loop, where our "prenoetic" actions serve to materialize an ongoing cognitive process, a process which finds itself reinforced by the awareness of such actions. As such, gesture not only helps bring about an act of thought, but can also serve in the transition of one cognitive state to another.

In "Is Cognition Embedded or Extended," Michael Wheeler uses gesture to argue for his thesis of embedded cognition, a less radical approach than the EMH. Where extended cognition takes external environmental materials to be (possibly) legitimate realizations of the mind, embedded cognition holds that cognition is realized solely in the brain but can causally depend on the non-neural body and external devices. In dialogue with Clark's previous essay and Gallagher (2005), Wheeler argues that although gestures meet the criteria for cognitive self-stimulation, this does not cement the fact that such gestures are also the *material realizers* that instantiate cognitive states. Hearkening to a constant criticism of the EMH (Adams and Aizawa: 2008), Wheeler argues that though gestures may have a causal impact on cognition, they themselves do not constitute such cognition. Wheeler's essay is long and intricately argued, but a rewarding read for anyone unsure where she stands on the embedded versus extended debate. It is a remarkable fact keeping with this collection that the crux of Wheeler's position comes down to the hands.

The final essay in this section changes things up a bit, and we move from extended and embedded cognition to joint attention and symbolization. "Pointing Hand: Joint Attention and Embodied Symbols" by Massimiliano Cappuccio and Stephan Shepherd is a fascinating work with deep significance for anyone interested in anthropological and evolutionary-linguistic questions concerning hominization and glossogenesis. They frame their essay amidst the debate between dispositional and representational theories of social cognition. The first implies that for joint attention to obtain, the involved parties must engage in "reciprocal coordination mediated by embodied intentions" (Cappuccio & Shepherd, 304). The latter account specifies that the respective parties must form reciprocal "theories of mind" through which they reciprocally infer relevant propositional states concerning the attentional target. It should come as no surprise considering the other essays in this collection that the authors find fault with the representational account. They go on to provide a novel account of the development of symbolic joint attention, which is or can be representationally mediated, through a study of pointing gestures. Their basic thesis is that declarative (informationally assertive) pointing occupies a central role in the move from basic joint attention (BJA) to a symbolically mediated joint attention (SJA). BJA could be simple gaze-following where there is a shared attention to a target, reciprocal attention between both parties, and an iterative awareness of the others' attention. The point is that it does not make use of representations. However, while necessitating the same conditions to obtain as in BJA, because of the addition of representations SJA has the added bonus of producing a rich and varied world of meaning that is not available within a paradigm of BJA. Such states of SJA occur commonly enough in the day to day world—people awaiting the turning of a stop light for instance. All that is necessary is that mutual attention is given to a target (T) whose features are intrinsically irrelevant "except instrumentally to make the recipients aware of some background information related to shared attitudes toward T" (Cappuccio & Shepherd: 307). In such cases it could very well be that the participants formulate theories of mind to infer each other's intentions regarding T. But Cappuccio and Shepherd argue that it is through gestures such as pointing that we arrive at the possibility of such advanced symbolization in the first place. The core of their argument is that declarative pointing *intrinsically produces* a primitive form of representational intelligence, one which phylogenetically and ontogenetically grounds further non-bodily symbolization. They claim that pointing "incarnates the possibility of communal attention (...) and is simultaneously recognized by all parties as explicitly produced to coordinate awareness" and thus "symbolizes awareness in a prototypical form" (ibid.: 305). As such, SJA arises long before we develop the capacity to infer peoples' mental states. The authors convincingly present their case, engaging with the work Tomasello, Racine, and Peacocke, among others.

5. Manipulation and the Mundane

This section begins with Susan A. J. Stuart's essay "Privileging Exploratory Hands: Prehension, Apprehension, Comprehension." Stuart claims that the hands are essential to orienting the subject in space and its establishing the ego-centric point of view. Her thesis is that the "failure in functional symmetry establishes our physical spatiality and provides us with a situated perspective on the world" (Stuart: 331). What this ends up meaning is that the hands ground our presence in a manipulatable world: insofar as they are "enantiomorphic" organs (qualitatively identical but topologically non-identical) imbued with an inherent directionality, the hands by nature reach away from ourselves, in potentially different directions. Stuart bases her argument on Kant's proof for the existence of absolute space, though modifies it to highlight the point that it is only through *prehensive* exploration that we achieve any *com-prehension* of space at all. Stuart continues her discussion of the relevance of the hands as the "orienting structures of self-referential anchoring," noting the role of anticipation in corporeal exploration (ibid.: 335). Stuart mentions repeatedly the "enkinaesthetic" role of the hands (that they facilitate a felt "withness" with other people, agents, and things) but unfortunately does not expand on how this obtains. Her account in the end establishes the ego-centricity of the subject, but only gestures to its inherent intersubjectivity.

In "The Encultured Hand," Richard Menary takes a step back from embedded and extended theories of cognition to argue that the hands facilitate the internalization of cognitive artifacts, thereby enculturating the mind. His argument proceeds in several steps. He first demonstrates that elsewhere in the animal kingdom, intentionality need not exist solely in the head but can be mediated by the environment. Using the example of cricket song in mating rituals, Menary shows that intentional directedness is structured in a triadic manner: a) a male cricket produces song b) a female cricket's dedicated interneurone hears the song and locates the male c) the crickets reproduce. Intentionality thus seems not simply to be a property of a mental state but is rather a complex where an organism directs itself by *acting*, an acting which is mediated by its environment and/or other agents. In like manner, Menary argues that human intentionality is mediated by our cultural environment. Certain practices, like writing or arithmetic, have the precise aim of expanding our cognitive capacities. Such practices are culturally bound and thus are normative. But it is Menary's point that once the practices are internalized, the rules are no longer referenced. Building off his previous (2007) work on cognitive integration, Menary proceeds to give an account of how it is that our brains can internalize such practices. After all, a cultural practice such as writing, he points out, is only 10,000 years old. There is obviously no gene for writing. How is it the case that our brains can internalize such a practice so well? His answer is that the Pleistocene brain is phenotypically plastic (a version of what Bernard Stiegler might call our "Epimethean default of origin") (Stiegler:

1998). Such plasticity evolved, Menary claims, as such as an adaptive response to environmental contingency. Menary goes on to show that learning-dependent plasticity "reformats the representational capacities of the brain in terms of public representational systems" (Menary: 357). The essay continues by exploring how such top-down cognitive internalization is coupled with more bottom-up cognitive niche construction, and closes by opening the social-political question concerning how our potentialities for action in the world fluctuate depending on what processes of integration we have undergone.

The section closes with the Zdravko Radman's "On the Displacement of Agency: The Mind Handmade." The essay pits itself against the idea that agency is derived from some sort of centralized cognitive executer. As Radman points out, such a conception is the offshoot of a more intellectualist picture of cognition in general, where an action must be the result of some propositionally representable mental activity. This picture is challenged not in favor of a picture of "blind embodiment and naïve coping" but rather with the alternative idea that there exists practical knowledge generated apart from abstract rationalization that still can effectively cope with the world (Radman: 370). For Radman, the hands are the perfect vehicle unto which we can ascribe something like bodily agency—they can skillfully and actively cope with their environment without recourse to central deliberation. The essay begins with a discussion of Gibsonian affordances where Radman illustrates that what is graspable to a subject depends on what matters to her in her ecological niche. It continues with an account of "manual perception" where it is argued that much of manual action depends on the interplay between manual guesswork and manual intelligence, or the idea that action is largely determined by "casting a net of the probable onto what will be picked out as our actuality" (Radman: 382). Radman cogently argues that agency is often displaced from the "inside our heads" to the points of engagement with the objective world (often the hands), and concludes that we should think of ourselves as a "participatory being that engages in worldly affairs (…) without having always to consciously deliberate" (Radman: 389). This conclusion is interesting but one wishes that Radman would have gone further. Theorists such as Lambros Malafouris (2008, 2013) and Colin Renfrew (2004) have argued that agency, and indeed, intentionality, should be thought of as co-operatively emergent phenomena not necessarily localizable within any single specific body. A thing in the world might have agentive status insofar as I am its possible "patient."[65] And as its patient I can realize an intention-in-action that such an

[65] The concept of patiency is developed by Alfred Gell (1998) for the purpose of illustrating the distributed and relational nature of causality with respect to agency. A patient is the counterpart to an active agent, though its passivity does not entail a lack of agentive power. A patient may in a different situation be an agent, and vice versa.

agent calls forth. One could thus view agency and intentionality as situationally emergent events in which things in the world fluidly participate depending on the interactive conditions present at a given place and time. Such a view surely seems compatible with an outlook that labels the human subject a "participatory being." But Radman's account only seems to displace agency from out of the head and into the hands. We are left to wonder whether he would endorse any further displacement.

6. Tomorrows Hands

In the last chapter in the book, "A Critical Review of Classical Computational Approaches to Cognitive Robotics: Case Study for Theories of Cognition?," Ettiene Roesch draws on his experience at the Brain Embodiment Laboratory at the University of Reading to elaborate contemporary robotics research. It may surprise some to discover (though not when one considers the themes of this volume) that attempts to design a robot with functional hands have been met with little success throughout the history of humanoid robot production. And if the hands are the gateway to the world, it is clear that contemporary research is not yet in a position to provide us with any robots with meaningful active relations. Roesch closes by maintaining that enactive robots do not yet exist. However, if one were to come about, it would have to comport itself not as a deliberative and precisely calculating machine, but as skillful and dynamic entity in constant adjustment with its environment.

To conclude, Radman has done us all a service in editing this volume—the first of its kind to my knowledge. As hopefully evidenced here, there is more than enough material in this book to fuel numerous discussions down the road. And with such a high frequency of conceptual interplay between its various essays, one should not be surprised to flip open a journal and happen upon replies and rebuttals to and from its various authors. It's easy to see how such a volume provides the opportunity for further nuance and specialization. But it's also a treat to come in at the ground floor of such an enterprise.

References

Adams, F., and K. Aizawa. 2008. *The Bounds of Cognition*. Malden, MA: Blackwell.

Gallagher, S. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gell, A. 1998. *Art and Agency: An Anthropological Theory*. New York: Oxford University Press.

Iverson, J. and E. Thelen. 1999. Hand, mouth, brain: the dynamic emergence of speech and gesture. *Journal of Consciousness Studies* 6:16-40.

Malafouris, L. 2013. *How Things Shape the Mind*. Cambridge: MIT Press.

Malafouris, L. 2008. At the potter's wheel: an argument for material agency. C. Knappett & L. Malafouris, eds. *Material Agency: Towards a Non-Anthropocentric Approach*: 19-36. New York: Springer.

Marx, K. 1992. Economic and philosophical manuscripts. R. Livingstone and G. Benton, eds. *Early Writings*: 279-400. London: Penguin.

McNeill, D. 2006. *Gesture and thought*. Chicago: University of Chicago Press.

Menary, R. 2010. Introduction to the special issue on 4E cognition. *Phenomenology and the Cognitive Sciences* 9:459-463.

Menary, R. 2007. *Cognitive Integration*. New York: Palgrave MacMillan.

Piaget, J. 1952. *The Origins of Intelligence in the Child*. London: Routledge & Kegan Paul.

Piaget, J. 1954. *The Construction of Reality in the Child*. London: Routledge & Kegan Paul.

Renfrew, C. 2004. Towards a theory of material engagement. E. DeMarrais et al., eds. *Rethinking Materiality*: 23-31. Cambridge: McDonald Institute for Archeological Research.

Spelke, E.S., K. Breinlinger, J. Macomber, and K. Jacobson. 1992. Origins of knowledge. *Psychological Review* 99:605-632.

Stiegler, B. 1998.*Technics and Time, 1. The Fault of Epimetheus*, trans. R. Beardsworth. Stanford: Stanford University Press.


artist of the issue

AVANT, ISSN: 2082-6710
avant.edu.pl/en
OPEN ACCESS

Still life like a sort of motion

Interview with Robert Lemay[66]

by Monika Włudzik & Witold Wachowski

The interview was realized in August 2014.

You live in a special place in the world. You wrote on your blog that you wouldn't appreciate flowers and summer as much if it were not for the seven months of winter in Edmonton. Do you find wild roses, the floral emblem of Alberta, in any way inspiring? What other places do you enjoy? Botanically or otherwise?

There are wild roses growing on a path near my house. A few years ago, I clipped a couple of them and took photographs but I never painted them. They had interesting textures and shapes which you don't get from greenhouse flowers but in the end they just didn't speak to me. For me to paint flowers there has to be a confluence of light, shape, and my own mood.

Alberta is the accident of my birth, but places I have chosen to visit usually have great museums. New York and Amsterdam are favorites for their streetscapes and friendly people.

What techniques do you use if you can share the secret? What are your expectations from paints, paintbrushes and lighting while you work?

My techniques are fairly direct and come from observation of the painters whom I admire. For many years Vermeer and Velasquez were top of my list but lately I've returned to Rembrandt, de La Tour, and Degas as my teachers. Part of the beauty of being a painter is one can look over the shoulder of these long dead artists and learn directly how each one shaded drapery or painted highlights. Sometimes from reproductions, but most often from viewing originals. This is why I try to travel.

My basic technique is wet into wet where I start with a focal point and try to match the color and tone from the photograph. I then adjust and work out-

[66] The Artist's webpage ***http://robertlemay.com***

wards in facets in a way I think Cezanne would have. So, not the classical roughed-in under painting, but a more direct approach. I'm trying to draw slowly and then paint more freely once I've got the shapes I want.

I'm primarily self-taught in this, because at the University of Alberta in the early eighties, it was a school which was primarily influenced by Colour Field painting. So we had Kenneth Noland, Antony Caro, Stanley Boxer and Clement Greenberg as visitors. All important in their field, but not really useful for someone who wants to paint figuratively.

My expectations from paint and brushes is the hope that they transform a working photograph into a painting which is not merely a record but an object with its own truth and coherence as an experience.

It seems that you know all about the properties and secrets of colours and light. Are any aspects of your work that you see as particularly intriguing or, perhaps, challenging?

Art since the Renaissance has used the idea of the canvas as a two-dimensional surface containing the illusion of the third dimension. The secret of the illusion is the tonal variation which suggests objects and space within the illusory world behind the picture plane. The challenge is to understand that each new generation reinvents the relationships and forms used—the sign system of painting. So that just because I use the Renaissance and Baroque as my guide, I don't paint old men in funny hats. The influences, both conscious and unconscious, of the time one finds oneself in is the intriguing part of being an artist.

What is so special about flowers? How do you see flowers? Are they your signature mark, like bottles for Morandi?

There was a time when I thought I would never paint another flower. Other subjects, like fruit, drapery, books, typewriters and empty soda cans, seemed to address both formal and emotive concerns for me more than flowers. But then I saw a new way to explore the significant form of flowers. The enlarged scale and simplified presentation eliminated other aspects of the still life—the table plane, space between objects, etc, which I had grown tired of.

Painting a flower five feet tall transforms it for the viewer. This is why most traditional flower painting doesn't interest me. The scale of my flowers relates more to figure painting.

What are your views on botanical art? Do you find Victorian botanical plates inspiring? You seem to like the semi-scientific properties of the white background effect. Do you have your favourite botanical artists? How important are illustrative concerns in your work?

I like the quality of a lot of early botanical art. They were working alongside scientists with the same goal—to look closely and understand what they were seeing—not an idealized "art" viewpoint, but the same desire to observe and record. I currently use a light, uninflected background to highlight the observation of form in the flowers. Today, photographs can probably illustrate and document better than paintings. The painting is, instead, a record of the artist's interaction with the subject.

[Photo from Robert Lemay's archive]

Have ever collaborated with any conservationists or horticulturists?

I've read a little about flowers—the symbolism of the rose, why the Chinese like plum blossoms, but this isn't necessarily how I "know" flowers. My interest is in what I can do with the "ready-made" shapes and colours of flowers, trying to extend the still life tradition in this way.

You told the editor of *Edmonton Journal* that you have always painted still lifes and never tried other genres. Have you never tried to paint any objects in motion?

Someone once said, "still life is about contemplation, and therefore not for the young." I've always been "old" and my most special moments have been in quiet observation and reflection. Working from photos, I look at a split second in time, an instant, and then expand it in the painting of the work over the course of weeks. This is a sort of motion.

You are seen as a realist painter. It seems that your objects have their own, profound individualities, like human faces, beautiful and true, with their wrinkles and bruises. Is this anti-perfection stance your deliberate intervention in our photoshopped reality?

The realist impulse in art is always to see what is there, what actually is. Caravaggio and Vermeer show us aspects of this. Marc Quinn has made some paintings based on his sculptures of flowers frozen in perfection. My flowers try to show the life cycle and connect to the memento mori tradition.

There is also an abstract quality to your paintings. We think that it is possible to see your magnified flowers as reminiscent of imposing architectural forms or Renaissance draperies and clothing, as their ordinariness is diminished by their size. How would you define your position in relation to abstraction or, for that matter, abstract art?

I like all kinds of art. Cy Twombly is one of my favorite painters. Marina Abramović's work is intriguing. But I seem most drawn to the depiction of "things." Most art is a form of abstraction. You are taking a part to represent a whole but many have even questioned whether there is a truly non-representational art. Robert Motherwell said there is no such thing as a non-objective red, that red only had power as a pigment because of its associations.

In terms of mood, your paintings seem to evoke quite contemplative, zen-like musings on the transitoriness of nature. Are your paintings in any way philosophical?

I think all paintings have to ask a question. Since they are not for depicting or recording great battles or kings any more they must serve another purpose. Zen and Wabi-Sabi spiritual values state that Truth comes from the observation of nature and that greatness exists in the inconspicuous and overlooked object.

Two rather typical questions: How was your own artistic taste shaped, starting, let's say, from early childhood? Are there any old masters who have had a decisive influence on your style and technique?

People have observed that infants' earliest impulses are to reach out and grab whatever objects are within their grasp. This may be the original impulse to still life—the exploration of the visual and tactile nature of objects outside of ourselves.

As a kid, I always loved history. The school library had books about ancient Britain, Rome, Egypt. Partly, I loved escaping into the past, but I also liked finding out about other people and cultures. So as part of general history, I started reading about Leonardo, Michelangelo, Rembrandt, and they, not Napoleon, became my heroes.

Leonardo still influences me because it was the Renaissance which began the scientific inquiry into vision, geography, the nature of the cosmos, mapping, lenses, the exploitation of natural resources, also known as "the voyages of discovery," modern banking, etc. In short, the processes we are still experiencing now. So my work is based on the optics and perspective of the Renaissance.

Would you say that your temperament and personality is expressed in your paintings?

Yes. Matisse said the work is finished when the artist has painted themselves out of it. But the impulse to make art must derive from the need to express one's own subjective thoughts and feelings about the time and place one finds oneself in.

How much, in your opinion, have the processes of art perception and appreciation changed since, for example, the Renaissance? To what degree are they now different processes?

My central idea is that the world we now live in owes a lot to the mapping and classification and quantifying of the natural world begun in the Renaissance. One of the biggest changes in the dissemination of art has been the internet. We wouldn't be conversing without it. And this has made it easy to see what's going on in Hong Kong, New York, or Warsaw, without leaving home, or having to read ARTnews Magazine. As with music and books, there are fewer gatekeepers deciding what people can see or hear.

To what extent viewers are present in your art? What do you demand from them? Would you say that your paintings have quite a specific or rather a more universal appeal?

In still life, the viewer sees the same thing the artist sees: the object staring back. This is what I want to communicate to the viewer—the excitement I feel when looking. When you paint non-referentially you are asking the viewer to "speak your language." The beauty of realism is you can manipulate the abstract qualities of art—shape, colours, space, light—using recognizable objects, everyday objects. Working painters understand that a realist painting and an abstract painting have more in common as mark making on a flat surface than either has with the "real world." Any artist's appeal—be they writers, musicians, or painters—consists of the number of people who feel moved by the work. This may be 10 or 10 million.

Robert Lemay

> *This was almost the last rose in our garden in the fall and it had really begun to wither. Roses that are outside change colours in interesting ways because of the sunlight and this one bloomed pink and peach, but changed to cream, purple and brown as it aged. I photographed it from many different angles but chose the vertical pose because the flower almost looked like a dancer with a head, arms and a skirt at the bottom.*
>
> [R. Lemay[67]]

Robert Lemay graduated from the University of Alberta with a BFA in 1984. He began exhibiting in galleries the following year, in 1985. He has had over 20 solo shows in cities across Canada, including Calgary, Vancouver, Edmonton, Toronto and Montreal. He also shows his work in Santa Fe, New Mexico. Lemay has travelled extensively, studying the art in museums in Europe and the United States. He has visited New York, Los Angeles, Germany, Italy, Greece, Portugal, and Spain. R. Lemay lives and paints fulltime in Edmonton with his partner, Shawna Lemay, a writer, and their daughter, Chloe.[68]

AWARDS

2007 Alberta Foundation for the Arts Projects Grant
1996 Alberta Foundation for the Arts Projects Grant
1990 Elizabeth Greenshields Foundation Grant
1989 Elizabeth Greenshields Foundation Grant

LECTURES

2007 Alberta Society of Artists, September 13, 2007

[67] Source: *http://robertlemay.blogspot.com/2012/11/full-blossomed-magnificence.html*

[68] Source: *http://www.wallacegalleries.com/artists/robert-lemay*

COLLECTIONS

Alberta Art Foundation,; Alberta Energy; Alberta Treasury Branch; Bennet, Jones, Verchere; Canada Council Art Bank; Canadian Embassy: Beijing, China; Canberra, Australia; Chevron Oil; Cliff Lede Wineries, Yountville, California; Edmonton City Hall; Foreign Affairs; Graceland University (Lamoni, Iowa); Interprovincial Pipeline; Lethbridge Regional Hospital; Norcen Resources; Ocelot Energy; Shaw Communications; Suncor Energy; Westin Hotels

SOLO EXHIBITIONS

2013	"Flower," Douglas Udell Gallery, Edmonton
2011	Douglas Udell Gallery, Edmonton
2010	"In a Narrow Space," Shayne Gallery, Montreal
2009	Wallace Galleries, Calgary
2008	"Calm Things," Douglas Udell Gallery, Edmonton.
2007	Shayne Gallery, Montreal
2006	Wallace Galleries, Calgary
2005	Douglas Udell Gallery, Edmonton
2004	Shayne Gallery, Montreal
2003	Douglas Udell Gallery, Edmonton
2002	Shayne Gallery, Montreal
	Wallace Galleries, Calgary
2001	Douglas Udell Gallery, Edmonton
2000	Wallace Galleries, Calgary
1998	Hollander York Gallery, Toronto
	Douglas Udell Gallery, Vancouver
1997	Wallace Galleries, Calgary
1996	Douglas Udell Gallery, Edmonton
1995	Hollander York Gallery, Toronto
	Wallace Galleries, Calgary
1994	Douglas Udell Gallery, Vancouver
	Wallace Galleries, Calgary
	Hollander York Gallery, Toronto
1993	Douglas Udell Gallery, Edmonton
1992	Woltjen/Udell, Edmonton
1991	Wallace Galleries, Calgary
1989	Woltjen/Udell, Edmonton
1988	Woltjen/Udell, Vancouver
1987	Woltjen/Udell, Edmonton

GROUP EXHIBITIONS (SELECTED)

2011 "For the Love of Art," Wallace Galleries, Calgary, Alberta

1999 "New Talent Invitational" Denise Bibro Fine Arts, New York City
"Alberta: Six Degrees of Separation," Prairie Regional Art Gallery
Grand Prairie, AB

1998 "Multiple Realities," Muttart Public Art Gallery, Calgary

1997 "Still Life," McMullen Art Gallery, Edmonton

BIBLIOGRAPHY

2010 Canadian Interviews. "The Moment of Astonishment: Interview with Robert Lemay." canadianinterviews.com.

2010 Ryan, Janice. "Oil Paintings Animate Common Objects," Edmonton Journal, Friday, June 17, 2011.

2009 Notebook Magazine. V.3, Issue 8. "Shawna & Robert Lemay: Calm Things."

2008 Bouchard, Gilbert. "Lemay Reinterprets Iconic Art in his Neutral Realist Style," Edmonton Journal, Friday, April 11, 2008.

2006 Bouchard, Gilbert. "Alberta: What is Visible," Galleries West Magazine.

2005 Bouchard, Gilbert. "Lemay Breathes New Life into Form", Edmonton Journal. Oct 7, G8.

2005 Lemay, Shawna. "Precarious", Artichoke. Summer, 2005, pp. 38-45.

2003 Kellog, Alan. "Catching Up With the Beautiful People", Edmonton Journal.

2001 Bouchard, Gilbert. "Real Life Frozen in Still-Life Mastery", Edmonton Journal. May 4, E14

1999 Gustafson, Paula. "Robert Lemay at Douglas Udell Gallery," Asian Art News, March/April 1999

1998 Image reproduced. Alberta Views, Spring Vol. 1, No.2, pp. 15.

1996 Mandel, Charles. "Still-Life Showing so Conventional it's Uncongenial", Edmonton Journal. January 25, C5.

1993 "Time is Frozen in a Painter's Still Life Studies", Edmonton Journal. December.

1992 "Artist Follows Cezanne's Example", Edmonton Journal.

1989 Beauchamp, Elizabeth. "Summer Candid Pictures Can Lift Winter Mood", Edmonton Journal. December 2, D4.[69]

[69] Source: *http://robertlemay.com*

Contributors to this Issue

Kenneth Aizawa

Kenneth Aizawa is Professor of Philosophy at Rutgers University. He is a philosopher of psychology. In recent years, he has been working in two principal areas. First, he has been writing about the hypothesis of extended cognition according to which cognitive processes sometimes extend from the brain into the body and external world. Second, he has been collaborating with his friend Carl Gillett on the multiple realization of psychological properties. Books: *Building the Brain* (with C. Gillett), in progress; *The bounds of cognition* (with F. Adams), 2008, Blackwell Pub; *The Systematicity Arguments* (series *Studies in Brain and Mind*), 2003, Boston.

https://sites.google.com/site/kenaizawa

Krystyna Bielecka

Krystyna Bielecka is a Graduate Student of Philosophy at Department of Epistemology, Institute of Philosophy, University of Warsaw. Her main research interests include the concept of representation in philosophy of mind and cognitive science. The role of mental misrepresentation in the modern theories of representation is the focus of her doctoral thesis. She is a member of the European Network for the Advancement of Artificial Cognitive Systems, Interaction and Robotics (EuCogIII), as well as the Centre for Philosophical Research.

https://uw.academia.edu/KrystynaBielecka

Matthew Bower

Matt Bower's research interests are in phenomenology and philosophy of mind/cognitive science. He received his PhD from the University of Memphis, with a dissertation working out issues in the genetic phenomenology of Edmund Husserl related to perception. Currently a Visiting Lecturer at the University of Central Florida, Matt is interested in the relevance of genetic phenomenology to studies of cognitive development, the function of affect in perception, and the issue of perceptual content/representation.

https://ucf.academia.edu/MattBower

Robert Briscoe

Robert Briscoe is Associate Professor at Department of Philosophy at Ohio University. His research takes an empirically oriented, interdisciplinary approach toward a range of topics in the philosophy of cognitive science, philosophy of perception, and philosophy of mind. Topics he has written about include the role of action in perception, spatial representation, the dual systems model of visual processing, mental imagery, cognitive penetration, sensory substitution, depiction, and pictorial experience.

https://sites.google.com/site/roberteamonbriscoe

Fred Cummins

Fred Cummins is Senior Lecturer at School of Computer Science & Informatics at University College Dublin. His research interests: Joint speech as found in prayer and protest; Post-cognitive approaches to the foundations of cognition; Temporal patterning in speech production and perception; Speech rhythm; Dynamic modeling within cognitive science; Gesture, Gaze and Blinking; Speech rate; Conversational interaction; Individual and social cognition; Collective experience; Relation between latter-day approaches to Cognitive Science and recurrent themes in Eastern Religious Philosophy. Co-editor of *Multidisciplinary Aspects of Time and Time Perception*, 2011, Berlin.

https://www.csi.ucd.ie/users/fred-cummins

Christopher Drain

Christopher Drain holds degrees from U.C. Berkeley and Boston College, and is completing his PhD in philosophy at Villanova University. He is interested in social ontology and the philosophy of technology, broadly construed, as well as the history of philosophy. He lives in Philadelphia, Pennsylvania.

https://villanova.academia.edu/ChrisDrain

Ralph D. Ellis

Ralph Ellis received his PhD in philosophy at Duquesne University and a postdoctoral M.S. in Public Affairs at Georgia State University. He has worked as a social worker as well as teaching philosophy, and is interested in applied phenomenology and integrating the social sciences with philosophy of mind. His books include, among others, *An Ontology of Consciousness* (1986), *Questioning Consciousness* (1995), *The Caldron of Consciousness: Affect, Motivation, and Self-Organization* (2000), *Love and the* Abyss (2004), *Curious Emotions* (2005), *Foundations of Civic Engagement* (2006, co-authored with Jim Sauer and Norm Fischer), *How the Mind Uses the* Brain (2010, co-authored with Natika Newton), and a critical thinking textbook, *The Craft of Thinking*. Ellis is also co-editor with Peter Zachar of a book series, *Consciousness & Emotion* *www.benjamins.nl/jbp*

Shaun Gallagher

Shaun Gallagher is the Lillian and Morrie Moss Professor of Excellence in Philosophy at the University of Memphis. His areas of research include phenomenology and the cognitive sciences, especially topics related to embodiment, self, agency and intersubjectivity, hermeneutics, and the philosophy of time. He has secondary research appointments at the University of Hertfordshire (UK) and the University of Wollongong (AU). He is editor of *Phenomenology and the Cognitive Sciences*. Selected books: *Phenomenology*, 2012, London; *The Phenomenological Mind* (with D. Zahavi), 2008 & 2012, London; *Brainstorming: Views and Interviews on the Mind,* 2008, Exeter; *How the Body Shapes the Mind,* 2005, Oxford.

http://www.ummoss.org

Patricia Grosse

Patricia Grosse is a graduate student (PhD Candidate almost) in Villanova University's Philosophy Graduate Program. Her dissertation entitled "Embodied Love and Extended Desire: An Undertaking of Embodiment" draws together Feminist Epistemology, Extended Mind, and Late Antique Philosophy and Theology in order to develop a deeper account of human being-in-the-world. She lives in Philadelphia, Pennsylvania with her two cats.

https://villanova.academia.edu/PatriciaGrosse

Tomasz Komendziński

Tomasz Komendziński (PhD) is a researcher at the Department of Cognitive Science and Epistemology and at the Centre for Modern Interdisciplinary Technologies at Nicolaus Copernicus University. His main area of interest covers integrated theory of communication, embodied cognition and extended mind, neurophenomenology, enactivism, theory of interdisciplinary studies, disorders of consciousness and neurorehabilitation. He is editor of *Theoria et Historia Scientiarum. An International Journal For Interdisciplinary Studies* and author of two books.

https://umk.academia.edu/TomaszKomendzinski

Robert Lemay

Robert Lemay graduated from the University of Alberta with a BFA. He began exhibiting in galleries in 1985. He has had over 20 solo shows in cities across Canada, including Calgary, Vancouver, Edmonton, Toronto and Montreal. He also shows his work in Santa Fe, New Mexico. Lemay has travelled extensively, studying the art in museums in Europe and the United States. He lives and paints fulltime in Edmonton.

http://robertlemay.com

Qing Li

Qing Li is a full professor at the Department of Educational Technology and Literacies in Towson University. She has taught various courses in the fields of educational technology, game based learning, and mathematics education at the undergraduate and graduate level. Current research interests: Game-based learning; Enactivism; E-learning, technology integration in education; Mathematics education; Cyberbullying. Her most recent publications include *Cyberbullying in the Global Playground: Research from International Perspectives* (co-editor), 2012, London.

http://qbook.towson.edu

Przemysław Nowakowski

Przemysław Nowakowski is a graduate in Philosophy at Faculty of Humanities of Nicolaus Copernicus University in Torun. The title of his PhD thesis is "The body space as a basis for the embodiment of cognition. In search of cognitive conception of the body". Research interests: embodied and predictive theories of cognition.

https://independent.academia.edu/Przemys%C5%82awNowakowski

Jean-Luc Petit

Jean-Luc Petit is Professor of Philosophy at Université de Strasbourg. He is a French philosopher. He works in the field of phenomenology (the science of experience as lived by an 'I'), a field founded at the beginning of the 20th century by the German philosopher Edmund Husserl. He owes his introduction to the thinking of Husserl to his teacher and friend, Paul Ricœur. His publications include: *Complexité-Simplexité* (co-editor), 2014, Paris; *The Physiology and Phenomenology of Action* (with A. Berthoz), 2008, Oxford; *Les neurosciences et la philosophie de l'action* (editor), 1997, Vrin.

http://www.jeanluc-petit.com

David A Reid

David A. Reid is a full professor and director of the School of Education at Acadia University. He is now working at the Universität Bremen. He has been using enactivism as a theory of learning and a methodology in his mathematics education research for two decades. His other research interests include proof and proving, sociological perspectives on the emergence of disparity in mathematics achievement, and comparative perspectives on mathematics pedagogy. Author of *Proof in mathematics education: Research, learning and teaching* (with C. Knipping), 2010, Rotterdam.

http://www.acadiau.ca/~dreid

Robert D Rupert

Robert D Rupert is Professor of Philosophy at the University of Colorado, Boulder, as well as Professorial Fellow at University of Edinburgh. His longest running research interest is in naturalistic theories of mental content. He has a related interest in the connection between theories of mental content and the proper nonsemantic characterization of mental representations. He also writes about extended, embodied, and embedded cognition, the extended mind, group minds, mental causation, and various metaphysical issues related to the special sciences. Also, he wrotes a book about lots of these issues, *Cognitive Systems and the Extended Mind* (2009, Oxford).

http://spot.colorado.edu/~rupertr

Pierre Steiner

Pierre Steiner is Associate Professor of Philosophy and Cognitive Science at the Université de Technologie de Compiègne / Sorbonne (France), where he heads the research group "CRED" (Cognitive Research and Enaction Design). His domains of interest and publication include non-representationalism on mind and language (inferentialism); vehicle-externalism on cognitive processes & vehicleless approaches on conceptual mental episodes; enactive (autopoiesis-based) theories of cognition; intentionality and (social) normativity; pragmatism (old and new), and Wittgenstein. His forthcoming book, *Qu'est-ce que la pensée?* will be published by Vrin in 2015.

https://sites.google.com/site/pierresteinerweb

Ian Winchester

Ian Winchester is a professor at Faculty of Education at University of Calgary. His research interests include: Cognitive science; Theory/philosophy of higher education; Literacy; Science education; Theory of knowledge; Theory/philosophy of science; History of science; History of higher education. He is particularly interested in the question of what kind of an institution a university is and how can we make it better. His publications include *Understanding Literacy in Its Historical Contexts: Socio-Cultural History and the Legacy of Egil Johansson* (co-editor), 2010, Nordic Academic Press.

http://werklund.ucalgary.ca/educ_info/profiles/w-ian-s-winchester


GALERIA SZTUKI WOZOWNIA
ul. Ducha Św. 6, Toruń / tel. +48 56 622 63 39 / www.wozownia.pl
poniedziałek - nieczynne
wtorek – 12-20 - wstęp wolny
środa/czwartek/piątek – 11-18
sobota – 12 -18
niedziela – 12-18 – wstęp wolny
ART GALLERY
ul. Ducha Św. 6, Toruń / tel. +48 56 622 63 39 / www.wozownia.pl
WOZOWNIA
monday - closed
tuesday – 12-20 - free
wednesday/thursday/friday – 11-18
saturday – 12 -18
sunday – 12-18 – free

One of the most common questions in today's cognitive studies is the one regarding embodied cognition. The answer to this question draws our attention to many factors, including bodily actions, which also work to embody cognition. With this in mind, enactivism is included in discussions of embodiment. In the current issue we present texts in which a focus on enactivism itself is the leading topic. (...)

A certain common notional basis for enactivism is often pointed towards; it is comprised of such notions as autonomy, sense-making, structural-coupling, self-organisation, agency, action, and sensorimotor dependencies. It is still a long way from showing the relations between these notions and from unifying enactivisms into one common theoretical proposal in a satisfactory manner. However, not only does this not stop the researchers, but it also encourages them to further – especially critical – studies, which will allow enactivism to discover itself anew. (...)

[*Introduction*]