Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Philosophy, Science, Cognition, Semiotics (PSCS)

Ciclo XXX

Settore Concorsuale di afferenza: 11/C4

Settore Scientifico disciplinare: M-Fil/05

THE PROBLEM OF REPRESENTATION BETWEEN EXTENDED AND ENACTIVE APPROACHES TO COGNITION

Presentata da: Marta Caravà

Coordinatore Dottorato

Prof. Marco Beretta

Relatore

Prof. Claudio Paolucci

Esame finale anno 2018

Table of Contents

	Abstract	1
	Riassunto	2
	List of Figures	3
	List of Abbreviations	4
	Introduction and plan of the work	5
I	**The Mind and its varieties. Alternative Minds in the Cognitive Sciences**	9
I.1	Introduction. Approaching "4Es"	10
I.2	When we rise. From speculation to the "dirtiness" of experience	12
I.3	Moving the level of pertinence of semiotic systems. Towards an embodied account of meaning	23
I.4	Cognitive linguistics. Looking for an embodied matrix for meaning-acquisition	27
I.5	Grounded Cognition. Embodying modal representations	33
I.6	The body's βίος. Towards the very materiality of the body	42
I.7	Making sense of inter-action. Enactivist Embodiment	47
I.8	Embedded Cognition. Discussing the project of locating cognition in externalities	62
I.9	The Extended Mind. A radical theory to extend the mind?	71
I.10	Concluding remarks. Convergences and tensions in the "4Es Cognition" debate	83
II	**Activating Representations. Looking for the whole story**	88
II.1	Introduction. Approaching Action-Oriented Representations	89
II.2	Acting without representing? Practical knowledge and minimal representationalism	91
II.3	Cognitive experience is just being there...plus Action-Oriented Representations. Clark's approach to AORs	109
II.4	Representational criteria. Peircean interventions	120
II.5	Objections against Clark's AORs: Standing-for, Interpretability, Decouplability and the problem of activity in representation. Some Peircean suggestions	133
II.6	Enacting salience. Reassessing the enactivist proposal through Peirce's indexicality	148
II.7	Concluding remarks	164

III Predictive Processing. The problem of representations 167
III.1 Introduction. Approaching Predictive Coding 168
III.2 Setting out the problem: the theoretical continuity between AORs and Predictive Processing 170
III.3 An internalistic version of Predictive Coding: Hohwy's proposal 180
III.4 Clark's approach to anticipatory mechanisms: extending Predictive Processing? 190
III.5 Clark's Predictive Processing: representational flashbacks 201
III.6 The enactive approach to Predictive Processing: an alternative proposal 213
III.7 Concluding remarks 228

Conclusion and future work 230

Works cited 233
Web sources 255

Abstract

Recent works in philosophy of mind and the cognitive sciences draw an "unconventional" picture of cognitive processes and of the mind. Instead of conceiving of cognition as a process that takes place within the boundaries of the skull and the skin, some contemporary theories claim that cognition is a situated process that encompasses the human agent's boundaries. In particular, the Extended Mind Hypothesis (EMH) and the Enactive approach to cognition claim that embodied action is constitutive of cognitive processes, and thus of the mind.

Although both theories give an "extended" or "extensive" picture of cognition and of the mind, they disagree on the epistemic value of internal representations. The EMH claims that we need to posit internal action-oriented representations (AORs). AORs would account for action-selection, action-control, and for the prediction of incoming perceptual information.

The enactive approach to cognition argues against AORs. The concept of AOR does not fulfill the representational conditions necessary to talk about representations properly. Furthermore, AORs are expressive of an internalistic prejudice, which makes the EMH weak.

Moreover, a semiotic analysis of AORs shows that these items called "representations" are not active at all. Therefore, the epistemic posit of AOR plays no interesting job in the project of extending the mind in virtue of a reassessment of the concept of representation aimed at making it embodied and active.

Therefore I claim that the concept of AOR has to be rejected. Action-control, action-selection, and the anticipation of aspects of action-perception loops can be explained in a more enactive way. Embodied action in a field of affordances explains how agents respond selectively to environmental features and how action-perception loops are anticipated by the "affective agent".

Furthermore, the enactive approach to cognition - especially if coupled with a semiotic description of cognitive niches and with some insights from the affective sciences (e.g. appraisal of core relational themes) - gives an explanation of action that, in contrast to the EMH, is actually able to "extend the mind".

Riassunto

Alcune ricerche contemporanee in filosofia della mente e scienze cognitive dipingono un'immagine della mente non convenzionale. Invece di concepire i processi cognitivi come qualcosa che accade nella testa, esse sostengono che la cognizione è un processo situato, che oltrepassa i confini dell'agente cognitivo umano.

In particolare, l'Ipotesi della "Mente Estesa" e l'approccio enattivo alla cognizione sostengono che le azioni dei soggetti incarnati costituiscono i processi cognitivi, quindi anche la mente.

Nonostante entrambe le teorie descrivano la mente e la cognizione come "estese", c'è disaccordo circa il valore epistemico della nozione di rappresentazione interna. L'Ipotesi della "Mente Estesa" sostiene che la nostra teoria debba postulare rappresentazioni interne orientate all'azione. Queste spiegherebbero come il sistema cognitivo seleziona l'azione da performare e la controlla, e come esso anticipa l'informazione fornita dalla percezione. L'approccio enattivo alla cognizione argomenta contro le rappresentazioni interne orientate all'azione. Il concetto di rappresentazione orientata all'azione non rispetta i criteri secondo i quali si può parlare appropriatamente di rappresentazione. Inoltre, il concetto di rappresentazione orientata all'azione nasconde un pregiudizio internalista, che fa dell'Ipotesi della "Mente Estesa" una proposta debole.

In aggiunta, considerazioni di natura semiotica sul concetto di rappresentazione orientata all'azione dimostrano che queste rappresentazioni non sono davvero concepite come processi attivi. Quindi il concetto di rappresentazione interna orientata all'azione non sembra essere utile nel progetto filosofico di una "Mente Estesa", basato su una rielaborazione della nozione di rappresentazione in termini incarnati e attivi.

Sostengo allora che il concetto di rappresentazione orientata all'azione debba essere abbandonato. Il controllo e la selezione dell'azione, così come gli aspetti anticipatori di processi cognitivi basati sulla circolarità di azione e percezione, possono essere spiegati facendo riferimento alla letteratura sull'approccio enattivo alla cognizione. L'azione del soggetto incarnato, performata in un ambiente fatto di affordance, spiega sia come il soggetto risponde agli stimoli ambientali in maniera selettiva, sia come i processi cognitivi basati sulla circolarità di azione e percezione sono anticipati da un soggetto emotivamente tonalizzato.

Inoltre, l'approccio enattivo alla cognizione -specialmente se integrato con una descrizione semiotica delle nicchie cognitive e con alcuni concetti delle scienze dell'affettività, come quello di valutazione delle emozioni di base- fornisce una spiegazione dell'azione che, contrariamente all'Ipotesi della "Mente Estesa", è davvero in grado di "estendere la mente".

List of Figures

Figure 1	A-modal Symbol Systems	36
Figure 2	Perceptual Symbol Systems	38
Figure 3	Body, Action, Context, and Representation in second-wave Cognitive Sciences	87
Figure 4	Chains of Triads	142
Figure 5	Micro-level semiosis	145
Figure 6	Markov Blanket	182

List of Abbreviations

CP: Peirce, C. S. (1931-58), *Collected Papers of Charles S. Peirce*, Cambridge (MA), Harvard University Press.

EP: Peirce, C.S. (1992-98), *The Essential Peirce. Selected philosophical writings*, Bloomington, Indiana University Press.

MS: C.S. Peirce's Manuscripts, according to the classification given in R. S. Robin (1967) *Annotated Catalogue of the Papers of Charles S. Peirce, Cambridge*, Cambridge (MA), The University of Massachussets Press.

SS: Hardwick, C., S. (1977) [Ed.], *Semiotics and Significs: The correspondace between Charles S. Peirce and Victoria Lady Welby,* Bloomington, Indiana University Press.

W: Peirce, C. S. (1982-2009), *Writings of Charles S. Peirce. A Chronological Edition,* Bloomington, Indiana University Press.

Introduction and plan of the work

For the past twenty-five years, research in cognitive sciences have promoted a paradigm shift. First-wave cognitive science, sometimes called "Cartesian Cognitive Science" (Anderson 2003; Rowlands 2010; Wheeler 2005), identified cognition -and then the mind- with sub-personal computations on internal representations. The mind was said to stop at the boundaries of our skull.

More recent approaches to cognitive sciences and philosophy of mind known as 4Es (Embedded, Embodied, Extended, Enactive Cognition) question this claim. On the one hand, those theories claim that this classical paradigm for the study of the mind is not explanatory enough. The identification of the mind's boundaries with an internal representational software does not explain what agents do in everyday cognitive practices. Usually we perform cognitive practices in a given environment, and we do so by means of our bodies and of the exploitation of environmental resources. Therefore, to account for human agents' cognitive experience, cognitive sciences and philosophy of mind should take into account seriously the cognitive role of our bodies and of the environment we interact with and we act in. On the other hand, part of the contemporary research in cognitive sciences notices that the problem of classical theories consists in the use of the "wrong" notion of representation. Classical cognitive scientists' a-modal and proposition-like representations do not account for embodied and situated aspects of cognitive processes. Therefore, in order to be explicative, cognitive theories should rethink of the epistemic posit of representation in order to anchor it to the agent's cognitive experience.

Most part of the contemporary debate in the fields of cognitive sciences and the philosophy of mind deals with these two philosophical points: the role of representations in cognitive processes, and different ways to account for situated embodied action.

Some theories, such as "Grounded Cognition", some versions of "Embedded Cognition" and the "Extended Mind Hypothesis", try to explain the role of the body, of action, and of externalities in the mind's processes by rethinking of the concept of representation.

In my work I consider all those approaches to representations, although I focus on the representational approach promoted by the Extend Mind Hypothesis in particular. Indeed the Extended Mind Hypothesis is extremely

interesting in the debate because it explicitly endorses an ontological claim about the mind (Clark, Chalmers 1998; Clark 2008; Clark 2016): the mind can be realized outside of the skull.

Given the importance of internal representations in "Cartesian cognitive sciences", and the problems that the notion of representation entails, to understand whether the ontological claim of the Extended Mind Hypothesis is justified by the theory "at work", a detailed discussion of the way this theory deals with the problem of representation is required. Indeed, the Extended Mind Hypothesis tries to extend the mind by explaining external action –which is said to extend the mind- through the concept of Action-Oriented Representation.

My research questions are the following ones. Is this action-oriented representational proposal able to overcome the internal/external duality entailed by classical representational explanations of cognitive processes? Does the new concept of representation (i.e. action-oriented representation) developed in the Extended Mind literature play any interesting job in the philosophical project of an extended mind?

My working hypothesis is that the concept of action-oriented representation, instead of fostering an extended picture of the mind, hides an internalistic commitment. Therefore, I try to find other explanations to account for the extensive nature of the mind in the Enactive approach to cognition.

The Enactive approach to cognition claims that a successful explanation of the mind -able to account for cognitive processes as taking place in the "mindful" interaction between the cognitive agent and her environment- should get rid of any kind of internal representation. The explanation should rather focus on the very role of embodied action in a subjectively and affectively tuned world of affordances.

To develop my discussion about the role of representations in contemporary explanations of cognition and the mind, I will structure my work in three chapters.

In Chapter I, I will take into account the way the field of cognitive sciences changes over time. I will not offer a detailed excursus about all the theories of cognitive sciences and of the different kinds of representations they postulate. On the contrary, I will provide a short sum-up of the core ideas of classical cognitive sciences, and I will point out the centrality of representations in those explanations. Then I will consider different ways to reassess the concept of representation. Those proposals have been developed in order to solve the problems (i.e. the symbol grounding problem) the classic notion of

representation entails. In this discussion, I will focus on Cognitive Linguistics, Grounded Cognition, and Embedded Cognition, which I consider to be crucial attempts to "give a body" to representations and to account for the adaptive function of representational mechanisms. After this discussion, I will consider the Enactive approach to cognition and the Extended Mind Hypothesis.

This chapter is aimed at providing the theoretical background of the debate about representations that will be taken into account in the following chapter, and at emphasizing the differences among the theories of the contemporary debate of 4Es. In particular, the chapter will provide the conceptual means to distinguish the Extended Mind Hypothesis and the Enactive approach to cognition, my work deals with in detail. Both theories draw an "extended" or "extensive" picture of the mind, but they do that in different ways.

In Chapter II, those differences will be considered in depth. There I will deal with the problem of internal representation: this is one of the crucial points the Enactive approach to cognition and the Extended Mind Hypothesis disagree on. In particular, I will examine the concept of action-oriented representation. This one is pivotal in the explanation of practical knowledge offered by the Extended Mind Hypothesis. On the contrary, it is rejected by the Enactive approach to cognition, which shows the problems that this concept entails.

In my discussion about the concept of action-oriented representation I will consider the objections the Enactive approach to cognition makes against this epistemic posit. Moreover, I will provide more reasons to reject action-oriented representations by developing a semiotic analysis of this concept. This exam will show that action-oriented representations prevent the Extended Mind Hypothesis from giving a strong, "extended" picture of the mind. On the contrary, the enactive approach to cognition -especially if it is integrated with a semiotic perspective on eco-niches- promotes this explanation of the mind.

In Chapter III, I will deal with the problem of action-oriented representations again, but from another perspective. Indeed, in Chapter II, I will consider the very debate about action-oriented representations, postulated to account for action-perception loops and for the philosophical problem the explanation of those cognitive practices deals with (i.e. the frame problem). In Chapter III, I will consider the theory of Predictive Coding and some of the reassessments of this approach to cognition and the brain offered by the Extended Mind Hypothesis and by the Enactive approach to cognition. This chapter is aimed at understanding whether the Extended Mind Hypothesis' version of Predictive Coding (called "Action-Oriented Predictive Processing") is able to give a more successful explanation of action-perception loops than

that developed when it talks about action-oriented representations. Therefore, my discussion will focus on the concept of representation in the Predictive Processing framework, and on the way this approach to anticipatory aspects of action-perception loops works on Predictive Coding in order to make it fit with the Extended Mind Hypothesis.

My exam of this philosophical proposal will show that Predictive Processing still entails the same problems action-oriented representations-based explanations of action-perception loops entail: the theory is not suitable to "extend" the mind. Then I will deal with a different explanation of anticipation in cognitive processes. I will consider the "enactive version" of Predictive Coding, sometimes called "Predictive Engagement", and I will integrate it with some concepts from the affective sciences, in order to make some points of the explanation more clear.

Chapter I

The Mind and its varieties.
Alternative Minds in the Cognitive Sciences

I.1. Introduction. Approaching "4Es".

In this chapter, I begin to consider "4 Es" (Embodied, Embedded, Enacted, Extended) approaches to cognition. This discussion aims to understand the philosophical problems those theories seek to solve. Moreover, I consider how these approaches to cognition deal with some of the core concepts of the contemporary debate (i.e. the role of the body, action and representations in cognitive processes) in the cognitive sciences. This discussion is aimed at answering the following question: are '4Es' a solid philosophical coalition, or the projects pursued under the same umbrella follow philosophical paths that are the one at odds with the others?

To do that, in §I.2, I start my discussion considering the broad picture of "4 Es" approaches to cognitive sciences. I take into account the way those philosophical positions distinguish themselves from traditional cognitive sciences, and I describe 4Es' project as the attempt to go back to the "dirtiness" of cognitive experience.

In §I.3, I prepare the philosophical ground for the discussion that will take place in §I.4 and §I.5. In those paragraphs, I take into account Cognitive Linguistic and Grounded Cognition, in order to point out that new approaches to cognitive sciences born as a response to the "symbol grounding problem".

In §I.6, I consider an Embodied approach to cognition which differs from those presented in §I.4 and §I.5: "Biological Embodiment" (Gallagher 2011a). This passage is useful to develop my discussion for two reasons. First, it marks two distinct ways to conceive of Embodiment; second, it lays the foundations of one of the most discussed philosophical positions of the debate: the Enactive approach to Cognition. This one is examined in §I.7, where I consider some of the core ideas of Enactivism, and I try to explain how this approach to cognition draws a broad picture of the mind (similar *but* different from that suggested by the Extended Mind Hypothesis) developed in a non-representational fashion.

In §I.8, I discuss the philosophical stance known as "Embedded Cognition", in particular by taking into account the problem that I call the "location claim". This paragraph is also useful to distinguish different kinds of claims endorsed in the debate (e.g. dependence claim vs. constitution claim; ontic thesis vs. ontological thesis), and to draw some distinctions between the Enactive approach to cognition and the Extended Mind Hypothesis.

In §I.9, I take into account the Extended Mind Hypothesis in detail. I consider and answer to some objections against this approach to cognition and the mind. Moreover, I make the reasons of my critical position towards the

Extended Mind Hypothesis clear; I explain why I think that the hypothesis of the Extended Mind, as developed by Clark, is not the best way to think of the process of "mind-extension".

In conclusion, in §I.10, I discuss some differences and convergences about the way theories of the "4Es Cognition" debate consider the role played by the body, by action and by representations in cognitive processes. The discussion of these points is functional to explain why I claim that "4Es" do not constitute a robust philosophical coalition. Moreover, this paragraph is aimed at making clear how the Extended Mind Hypothesis and the Enactive approach to cognition, even if they point to the same direction (namely they understand the mind as dynamic, active and as encompassing the context in which cognitive practices take place) cannot be considered to be consistent the one with the other. In particular, a fundamental point of tension is considered. The Extended Mind Hypothesis, by endorsing the idea of cognitive impartiality (or that of complementarity) still gives an important role to internal representations in cognitive processes. This point is at odds with the Enactivist project, according to which, in order to get rid of any trace of internalism about the mind, we should reject internal representations (at least when we take into account low level cognitive processes, such as perception and motor coordination).

I.2 When we rise. From speculation to the "dirtiness" of experience.

An intuitive way to start manipulating the pieces of "4 Es" puzzle is to look at the background philosophical assumptions these different philosophical stances of the debate share.

On the one hand, the fact that Embodied, Embedded, Enacted, and Extended ("4Es") approaches to cognition and the mind belong from different philosophical traditions - from Merleau-Ponty's phenomenology (e.g. Varela, Thompson, Rosch 1991; Noë 2004) to analytic philosophy and AI (e.g. Clark 2008) – represents a great factor of differentiation. By recalling different philosophical traditions, theories within the "4 Es" cognition frame develop their argumentations independently the one from the others, also and especially disagreeing on some core theoretical points.

On the other hand, it should be noticed that, despite the great variety that characterizes this new field of studies, those theories have been said to constitute a "new science of the mind" (Rowlands 2010). They have been described as a new and revolutionary approach to the issues of cognition and mind that, despite many points of friction, apparently seems to constitute a sort of "theoretical coalition", which rises its voice against what has been the mainstream cognitive science for many years.

This second order look at "4E Cognitive Sciences", namely the idea that they should somehow be considered to be a "theoretical coalition" (at least for a provisional analysis), depends on the fact that all these new theories seem to break up with what the famous philosopher and historian of cognitive sciences Howard Gardner defines "the Mind's new science" (Gardner 1985), that is cognitive science as it was first established in the second half of the previous century. Hence, it seems to be the case that the philosophical ground "4 Es" share should be individuated in the breaking points with good old-fashioned cognitive science, in particular with some ideas that laid the foundations of cognitive science *as* a science, which those new approaches radically reconfigure or reject. Therefore, to approach 4Es' debate, it is worth to take into account the core points that determine their departure from classical cognitive science.

Early wave cognitive science -whose birth date is commonly individuated with the "Symposium of Information Theory" that took place at the Massachusetts Institute of Technology in 1956 (see Miller 2003: 142)- was settled as an interdisciplinary research field on cognition that gathered scholars

and scientists from philosophy, psychology, computer science, linguistics, neuroscience, and anthropology.

Despite this diversity, scholars of different research fields agreed on some fundamental assumptions (Gardner 1985: 38-49; Fusaroli, Paolucci 2011), which can be considered to be the pillars of the *cognitivist* epistemological (and methodological) paradigm. These fundamental assumptions can be summed-up in four points: *i) representations, ii) computers, iii) de-emphasis on affect, context, culture and history, iv) belief in interdisciplinary studies, and rootedness in classical philosophical problems.*

i) Representations.

Reacting to the behaviorist psychological paradigm, according to which we should study human intelligent capacities exclusively by considering the behavior that the agents display, first wave cognitive scientists adopt representations as fundamental epistemic posits.

It can be said that this conceptual move, rather than being motivated by an empirical necessity or by empirical studies, strictly depends on philosophical reasons, in particular on argumentative reasons. Indeed, this compelling necessity to think of mental representations as core features of "the new science of the mind" is an outcome of the criticism towards the argumentative structure of the behaviorists' theory[1], which is said to be unclear, vague and inconsistent. Behaviorism seems to entail a collapse of the *explanaundum* on the *explanans,* namely a circular argumentation that causes the inexplicability of the phenomena taken into account. When behaviorists try to explain intelligent behavior through the concept of behavior itself (described as a set of inputs and outputs data recordings), they end up with a misleading overlapping of causes and effects, which makes their causal explanation unsuccessful.

Then, it seems that epistemic entities such as mental symbols, compositional and syntactical rules, mental images, and so on are introduced within the cognitivists' explanatory framework for the sake of good argumentation, at first.

According to cognitivists, the postulation of this intermediate representational level between sensory stimuli (inputs) and behavior/actions (outputs) is an interesting and successful means to get rid of the circularity the behaviorists' explanation entails. Explaining cognitive functions by making use of representational concepts, a *locus* for cognition, different from that one individuated by behaviorists, is found.

[1] For accurate and detailed objections against the behaviorist explanatory paradigm see Chomsky's "Review of B. F. Skinner's *Verbal Behavior"* (Chomsky 1959).

Indeed, according to cognitivists, cognition not only depends on the manipulation of mental representations, but this intermediate, non-observable layer is considered to be the real *locus* of cognitive process. In this sense, sensory inputs can be said to be the first element of the causal chain of cognition, and behavior is the last one, the ultimate effect of a mechanism whose explanatory powers have to be individuated exclusively in the human head. Then, the collapse of *explananda* and *explanans*, and the collapse of causes and effects too, is avoided by identifying cognition with the "slice"[2] between inputs and outputs.

The "philosophical counterpart" of this general frame to study the mind is what is commonly known as "functional-computational" theory of mind.

Since this chapter is not aimed at giving a detailed theoretical and historical description of first wave cognitive sciences, I will not take into account the different theories that endorse this philosophical stance in detail. Nevertheless, since the functionalist-computational approach to cognition has been the mainstream explanation of the mind for many years, and given that great part of the debate concerning "4Es" Cognitive Sciences focuses on the reassessment or on the criticism of the functionalist stance, I briefly introduce the main points of this theory, whose implications for an explanation of mental phenomena will be made clear later.

According to the computational-functionalist approach to cognition, cognitive phenomena should be considered to be embedded in that language-like inner layer previously mentioned, whose items (representations) are manipulated by sub-personal computational operations, that is operations whose rules and passages can be explicitly individuated.

This is a formal way to reassess the assumptions of folk-psychological approaches to the mind, according to which the mind is a set of propositional attitudes or states (i.e. beliefs and desires), which cause intelligent behavior.

This is clearly exemplified in the way Fodor -the "complete cognitivist" (Gardener 1985: 81) sums-up the two core claims of his representational theory of the mind (RTM). The two claims are the following ones:

a) "For any organism O, and any attitude A toward the proposition P, there is a ('computational'/'functional') relation R and a mental representation MP such that
MP means that P, and

[2] The use of the word "slice" is not unintentional. Indeed, it recalls the phrase "sandwich model" coined by the philosopher Susan Hurley in her book *Consciousness in Action* (Hurley 1998: 401) to critically refer to the cognitivist paradigm.

O has A if O bears R to MP"

b) "Mental processes are causal sequences of tokenings of mental representations"
(Fodor 1987: 17)

Cognition takes place when there is an agent with a mental attitude (e.g. belief, desire) toward an intentional object (a representation), where mental attitude and intentional object are connected by a semantic relation. This is to say that a process is cognitive when it deals with mental states, whose contents are fragments of information represented by a mental symbol according to defined rules.

This means that, in order to explain why and how an agent displays intelligent behavior, the philosophical theory has to focus on the concepts of *state* and *information*. To explain *why* a subject is acting in a certain way in the defined context of a cognitive task, the theory should consider the state "activated" in the system at the moment "t" (the mental state of belief or desire, for example) and information that this mental state bears in virtue of its representational capacities.

For instance, thinking 'It's going to rain, so I'll go indoors' coincides with having a token of a mental representation whose meaning is 'I'll go indoors'. Meaning is caused, in a certain way, by the information contained in the tokened mental representation 'It's going to rain', that is a language-like translation of the agent's perceptions. The agent is able to behave in a certain way under certain conditions because she is able to sub-personally represent worldly stimuli, which become meaningful because they entertain causal and functional relations with other informational-representational states.

Then, it can be said that the functional relation that links information with dispositional states is meant to explain the "*how*" of cognitive processes, namely the way knowledge is acquired.

To explain the "how" of cognition within this complete cognitivist explanatory frame, the theory implicitly makes use of semiotic concepts. Indeed, the relation between information and cognition -that is the way information becomes meaningful within the context of holistic mental processes- depends on the notion of symbolic representation (Fodor 1987).

To briefly explain this point, I will make use of a terminology I am familiar with: the Peircean one. The choice to make use of this semiotic gloss- not very common in cognitive sciences - to clarify Fodor's concepts, is not only motivated by my philosophical background, in Semiotics and Pragmatism, but it is also motivated by a real theoretical problem. Indeed, there have been many

critiques of the way Fodor deals with the relation between mental signs and meaning/cognition[3], namely the "how" of cognition. I suspect that the lack of clarity of the RTM is not only a consequence of the confusion that characterizes folk psychology and its reassessment within a naturalistic frame (Descombes 1995: 84 – 89), but it also depends on the acritical use of the notion of symbol. Therefore, I will appeal to Peirce's concept of symbol to see if I can give a more intuitive and synthetic description of this point concerning Fodor's idea of processes of reasoning.

A symbol is a kind of sign that represents an object, where in Fodor's case that object is the content of a mental state, by exhibiting or displaying fragments of information as a token of a type. A symbolic representation is a general sign (CP 1.558) that stands for its object, denotes its object, according to a law (CP 2.249) that connects each representational instance to a representational type. This means that X is a symbol if it is lawfully connected to "X", where "X" can function as a law insofar its content is not derived from a previous representation.

Then, the idea the RTM seems to convey is that cognition consists in the manipulation of information, which becomes meaningful because it entertains a lawful relation with a set of general images (that have a propositional-like nature), which have an "original" content (Fodor 1987: 111-127; Adams, Aizawa 2010: 37).

Therefore, not only the "why" of cognition is individuated in a representational inner layer –because behavior is said to be *caused* by mental states, whose content is represented in a propositional-like structure- but also the "how" of cognition, namely the mechanisms through which knowledge is acquired, are explained by focusing on the way information acquires meaning when it is translated into a nomological representational relation.

The explanation of cognition given within this theoretical frame is fully representational: on the one hand, representations are the immediate/intentional objects of mental states, on the other hand they constitute the "immaterial" substrate of realizability of cognitive processes (Fodor 1987: 17).

ii) *Computers.*

Despite computers are not central in all cognitive scientists' everyday scientific practice (Gardner 1985: 46), the scientific production in the field of

[3] See, among others, Searle 1980, in which he develops the famous thought experiment of the Chinese room.

artificial intelligence deeply influences the way old-fashioned cognitive scientists work. Indeed, if one the one hand it can be said that AI researchers take human intelligence as a model, as something that it is supposed to be reproduced in a machine (Winston 1984), on the other hand cognitive scientists take artificial intelligence as a way to study human intelligence.

Cognitive scientists look at the way computers are built and at the way they display intelligent behavior as a material model to proof their theories. Artificial creatures are actors which perform (and in this performance test) the script (namely the theory) of cognitive scientists.

Moreover, computers play a significant role in the cognitivists' discursive practices, namely in the process where a piece of knowledge is developed and established as prominent at a given period.[4]

In fact, cognitive sciences' ultimate goal is giving a universal and naturalized account of the mind, which finds its place in RTM in virtue of the endorsement of a causal-informational theory of content. Therefore, computers are a good metaphor to think of cognition and the human mind: they are exactly designed to follow those causal, informational and representational laws at the core of mainstream theories of mind.

Moreover, computers are a "physical proof" of another pillar of functionalism, which is the multiple realizability of functional systems. Computer software can be realized by different kinds of hardware; in the same way, according to a functionalist perspective on the mind, mental phenomena can be individuated in virtue of their functional and causal relations, independently from the brain or the physical body of their realization.

Then, the computer metaphor is a successful way to "materialize" the encyclopedic knowledge[5] -namely knowledge that silently operates under the manifestation of a culture or subculture- that supports cognitivism as a leading scientific paradigm.

iii) De-Emphasis on Affect, Context, Culture, and History.

Mainstream cognitive sciences, although they think that affective, contextual, cultural and historical are parts of human lives, avoid making use of concepts that belong from those realms. This decision, which concerns the theoretical level of their scientific practices, is motivated by a methodological need. When cognitivism born, it was almost impossible to think of an experimental

[4] See Foucault, *The Archaeology of Knowledge* (Foucault 1969, in particular chapter II) for the use of the phrase "discursive practice".

[5] For a more detailed account of the relation between metaphor and encyclopedia see Eco's article "Metaphor, Dictionary, Encyclopedia" (Eco 1984).

paradigm able to take into account all of those variables at the same time and to keep them together in a unified model of cognition supposed to ground experimental practice. Moreover, this necessity of cleaning up their research from what can function as a factor of differentiation or as a factual constraint is motivated by an issue of practicality. According to cognitivists, if research had taken into account all those individualizing and phenomenalistic elements, cognitive science's project would have been emptied from its sense, because core of cognitive science's project was an explanation of universal mechanisms of human mind.

iv) Belief in interdisciplinary studies and rootedness in classical philosophical problems.

At the core of cognitive science's birth there was the attempt to reassess some classical philosophical problems within an interdisciplinary frame. Integrating knowledge and methods from different disciplines, cognitive science's core aim is finding solutions to the problem of cognition and the human mind, by trying to rephrase the questions that lead philosophical research for centuries: How does a cognitive agent look like? What is the best way to explain the relation between cognitive agents and their world? How does the mind acquire a grip on external reality? Which are the mechanisms cognition unfolds by? Which are the criteria we should endorse to define the boundaries of the mind?

Now, in virtue of this very short discussion of first wave cognitive sciences' core features, some considerations useful to understand why "4 Es" Cognitive Sciences embody an alternative theoretical stance can be offered.

It can be claimed that 4Es' objections towards classical explanations of cognition and the mind in cognitive sciences depend on the unsatisfactory way the cognitivist paradigm -and in general mainstream cognitive science- deals with some classical philosophical issues.

Indeed, what the cognitivist paradigm suggests is a Cartesian-like picture of the mind (Anderson 2003). This picture offers a fully internalistic and individualistic explanation of cognitive phenomena, which is grounded on the same set of conceptual oppositions critically assessed in centuries of philosophical research (especially by the phenomenological tradition and by the pragmatist one, I would say).

By endorsing a strong representationalist account of cognitive phenomena, namely by describing the mind as a computational-representational software, first-wave cognitive science draws a robust opposition between internal

realm/external realm, subject/object, cognition/action, mind/world, cognition/experience, mind/body, materiality/function, cognition/culture, cognition/affect, fact/value.

On the one hand, this set of oppositions could sound like a useful way to individuate accurate conditions to define cognition, and to set the boundaries of the mind. This would offer an explanation that, in principle, would be rigorous. On the other hand, exactly *because of* those same criteria, this approach to cognition seems to give an impoverished, flat and fragmented explanation of human cognitive life.

In particular, although those attempts to explain cognition give generalizable models of the mind that can be in principle applied to a large variety of cognitive phenomena – and this would guarantee a practical simplification of research practices- those models are not actually able to account of cognition in real life.

By claiming that cognition consists in a set of mental representations sub-personally computed, the cognitivist explanation seems not being able to have a grip on what human agents *experience* and *do* in their everyday life.

Behavior and action are put in the realm of out-puts, then they are not considered to be constitutive parts of cognitive processes. Rather, they are considered to be mere effects of a process entirely performed by an innate software embedded in human heads. Cultural, historical and affective features of human life are completely cut off from their experimental and theoretical framework. Therefore experiences that usually take place outside of the laboratory, namely experiences in which enculturated, affectively tuned embodied agents do something in their environment, are not taken into account as interesting subjects of study.

This suggests that the cognitivist paradigm cannot offer a cognitive model that intuitively justifies what humans do when they are engaged in their cognitive practices. It gives a "sanitized" explanation of cognitive phenomena that, precisely because of this sophisticated level abstraction, does not account for the *dirtiness* of human experiences. This makes cognitivism unable to fully develop cognitive sciences' project: explaining how human cognitive life unfolds.

"4Es" cognitive sciences' project can be seen as a revolutionary attempt to account for the philosophical problem of cognitive experience.

In order to account for this dirtiness of experience, for its folds and interstices, "4 Es Cognition" try to gradually reassess what is left outside of the boundaries of mind in the cognitivists' approach to cognition.

To briefly explain what will be taken into account in more detail later, it can be said that "4 Es" try to account for cognitive experiences by focusing on the following *dirty factors,* which are considered to be important *variables*, *constraints* or *constitutive* parts of cognitive processes:

i) The cognitive agent's body, considered in its neural components, as an anatomical body or as a living and moving body.
ii) The natural and cultural context where cognitive practices take place.
iii) The human agent's on-going interaction with the problems the cognitive situation displays. This interaction is said to be mediated by a skillful and embodied manipulation of the context or of parts of it (objects).

The progressive inclusion of those factors within the realm of the cognitive entails huge implications for a philosophy of the mind.

The argumentative structure of those explanations is similar to the cognitivst one: both explanations claim something about cognition in order to define or explain what the mind is. Nevertheless, the pictures of the mind these two approaches to cognition draw are completely different.

On the one hand, classical functionalist explanations of cognition, by endorsing a theoretical framework made of couples of oppositions, define the mind as something internal, completely distinguished from the world and the cognitive agent's body.

Moreover, by endorsing a mechanistic and representational perspective, shaped by the heuristic metaphor of the mind-computer, classical cognitivist explanations of cognition seem to draw a *passive* picture of the human mind. By considering actions and behaviors as mere effects of cognition and not as core ingredients of it, namely as unavoidable parts of cognitive processes, those explanations entail a strict dichotomy between theoretical knowledge and practical knowledge. Indeed, by the word "mind" they mean a set of mental states that have a propositional-like nature: having a mind means knowing that..., believing that..., desiring that... That is to say that the cognitivists' mind "does not care" about practical knowledge, about that "knowing how" or *phronetic*[6] knowledge, which is the ground of the encounter between the

[6] Here I use the phrase "phronetic knowledge" implicitly referring to the Heideggerian thought. This one, despite its great differences with research in cognitive sciences, is progressively becoming an important reference point to understand some issues in this field of study (Dreyfus 1972; Wrathall, Malpas 2000; Kiverstein, Wheeler 2012; Wheeler 2005). In

human agent and her world. They do not take into account what comes before any speculative attitude, what shapes our experience in its "average-everydayness", as Heidegger would say (Heidegger 1927: §5). To put it in other words, standard explanations of cognition and the mind do not consider cognition as something humans actively experience when they *do something* in their world, namely when they cope with those problematic situations that ask for practical solutions[7] we face all the time in everyday experience.

On the other hand, "4Es Cognition" -by including those said "dirty factors" in their explanations of cognition- redefine the mind as something that is not in principle separated by that same body, that same world the cognitivists' "Cartesian Mind" considers to be non-cognitive, and then non-mental. "4Es" respond to cognitivism's internalistic stance by prospecting a sort of "externalist turn".

As it will be shown later, not all those new theories endorse an externalist approach to the mind at the ontological level, claim according to which the mind actually encompasses what is "external" to the head. Nevertheless, they are said to endorse an externalist stance about cognition and the mind (Hurley 2010) at least at the epistemological level. Indeed, even when they explain cognitive processes by appealing to old-fashioned cognitive science's concepts (e.g. the concept of representation), they seriously consider the body and action in the environment, namely what the cognitivist explanation left unaccounted.

For those reasons, that is thanks to the endorsement of a progressively externalist and active account of cognition, those new approaches to cognition and the mind have been said to constitute a new, anti-Cartesian science of the mind (Rowlands 2010). This non-dualistic and anti-Cartesian framework is 4Es' common philosophical ground.

After this very short and general introduction to "4 Es", I start to explain what the different theories of the debate claim in more detail.

---

particular, I refer to the Heideggerian idea – prefigured in the seminal paper of 1922 "Phänomenologische Interpretationen zu Aristoteles", and then developed in *Being and Time* - that human's primitive encounter with the world is shaped and determined by the Aristotelian modality of *φρόνησις*, namely by a practical wisdom, a practical circumspection that guides human actions an behaviors in the environment without the need of a speculative look towards this same practical attitude.

[7] This way to explain practical knowledge echoes John Dewey's ideas described in *Logic. The Theory of Inquiry* (Dewey 1938). This reference could sound like not "well placed" in the context of cognitive sciences. On the contrary, as in Heidegger's case, it points to some theoretical directions that new studies in cognitive sciences are taking. See among others the quotation that opens Clark's book *Supersizing the Mind* (Clark 2008); Johnson 2010; Johnson, Rohrer 2007, Johnson 2006; Gallagher 2009; Menary 2016.

In particular, I will consider the argumentations concerning cognitive and mental externalism, and I will take into account different ways in which action plays a certain role in this multilayered process of "mind-extending".

I.3 Moving the level of pertinence of semiotic systems. Towards an embodied account of meaning.

The first critical response to the Cartesianism at the core of the cognitivist approach to cognition consists in thinking of a way to ground that symbolic, language-like system that, according to functionalism, constitutes the software of the mind.

As previously explained, the functional-computational approach to cognition describes cognition as taking place in the internal slice situated between inputs from the world and behavioral outputs. It explains the mind as something that is (in principle) isolated from the world and from the material substance that embodies the mind's representational software.

The semantics of that system of symbols is thought to be dependent on the internal relations that constitute the structure of the system in itself. Information that enters into the system through the sensory system is said to become meaningful because it is represented as a token of a type according to a compositional law, namely as an occurrence of a symbol-type, whose content is underived by previous representations.

Those approaches to cognition seem to endorse an explanatory perspective that has a peculiar semiotic nature.

I offer a citation from Hjelmslev's *Prolegomena to a Theory of Language* to make this claim clear.

> "[...] linguistic theory was established as immanent, with constancy, system and internal function as its sole aims, to the apparent cost of fluctuation and nuance, life and concrete physical and phenomenological reality. A temporary restriction of the field of vision, was the price that had to be paid to elicit from language itself its secret. But precisely through that immanent point of view and by virtue of it language itself returns the price that it demanded. [...] Instead of hindering transcendence, immanence has given it a new better basis; immanence and transcendence are joined in a higher unity on the basis of immanence" (Hjelmslev 1969: 127)

The fundamental idea of the linguistic method in its structuralist version- Hjelmslev and Saussure are representative of- consists in claiming that, in order to scientifically study linguistic phenomena as they are produced in the experience of a linguistic community, the scientist should operate a temporary abstraction from concrete, contingent, phenomenological manifestations of language. She should rather look at the structure that supports and motivates those phenomena.

In this semiotic tradition, structure is defined as a model built according certain operations of simplification, which allow the scientist to make different phenomena homogenous, looking at them through the lens of a defined point of view. For instance, I observe different human beings and, in order to find some common features that allow me to talk about different phenomena by appealing to homogenous conceptual tools, I produce a simplification. I can consider the human body as a web of relations I find in the skeleton, and I can draw a simplified graphic representation of this one. In this way, I individuate a structure all human beings share, a system of relations made of differences between elements that appear as discrete because they are modeled as occurrences of a topological, formal structure. Structure is not something that exists in itself; its ontological reality is undefined, or at least problematic.[8] It is rather a conceptual way to gather different items within a common frame (Eco 1968: 46; 48; 283- 288). Items that populate the structure acquire their conceptual reality in virtue of the oppositional relations that constitute the structure; the oppositional structure lights up those items as individual items, which can be conceptually manipulated. They show up as individual not because of their material, substantial or phenomenological features, but because the set of oppositions the structure is made of gives a formalization of a set of peculiar features (defined by the "direction" of the scientist's gaze), which show themselves as being different from another set of features.

This structure, defined by a differential topology, constitutes the plane of immanence semio-linguistic research should operate on. The concrete linguistic manifestations, conceived as a chaotic cluster of variables, transcend the semiotic analysis.

That is to say that, in order to develop a successful analysis of an object (language, in this case), the scientist is supposed to establish an accurate level of pertinence of her research practice, which defines the epistemic borders of the object of study. In the case of language, this level of pertinence is the relational, oppositional and componential structure of a closed or bounded semiotic system.

[8] There is a huge debate that deals with the problem of the ontological reality of structure. In my discussion, I endorsed Eco's position, according to which structure has a methodological value. Other scholars (Saussure, Jakobson, Tesnière, Hjelmslev, Piaget, Lévi-Strauss, Deleuze, Petitot), even if they do not actually make use of terms such as "realism" and "nominalism", endorse a different approach to structure, position that has been called "ontological realism". According to this approach, structure, in order to actually be an explicative concept, should be conceived as a *category* of possible experience (see Petitot 1985: 23-26). For a more detailed discussion of this point, see also Nöth 1990: 196:197.

By the expression "closed" or "bounded" semiotic system, I refer to a system in which each semiotic occurrence acquires its meaning in virtue of its lawful relation with determined exhaustive sets of classificatory alternatives, which are internal to this same system (Lemke 2000; Pattee 1995). Meaning never goes out, never escapes the system and never takes something from the grey and blurry landscape outside of this "semantic micro-universe" (Paolucci 2010: 49). The system is separated from the phenomenological, pragmatic and experiential landscape from a robust grating. This one is made of oppositions that follow a dyadic principle (for example contrariety, contradiction, and implication among couples of elements in Greimas, or oppositions between phonemes in Jakobson), or the law of "participatory oppositions", for example when an element of the structure takes part in the value of the opposite element, like what happens with extensive and intensive terms (Paolucci 2010:49- 79).

The functionalist explanation of cognitive phenomena seems to fit with this idea of closed semiotic system. According to this philosophical stance, cognition and meaning have to be studied as immanent to a well-defined level of pertinence, that is the formal level in which functional relations among semiotic mental entities unfold, abstracting from any kind of variable that cannot be modeled within this frame. Those variable are part of the non-scientific realm of transcendence.

In principle, there is nothing wrong with the attempt to scientifically study an object by modeling it through the concept of closed or bounded semiotic system. If the plane of immanence of the research is established by making use of pertinence as a methodological tool in a reasonable way - that is following the criteria of *observational adequacy,* according to which a semiotic model should be built taking into account its applicability to the empirical objects that it models (Hjelmslev 1969: 17)- a closed semiotic system can actually be explicative.

Indeed, by working on a formal schema that can be generalized to huge classes of objects, a closed semiotic system is, in principle, able to explain different kinds on phenomena in an economical way. It makes use of a minimal set of concepts to account for a variety of meaningful dynamics, variety that develops within the boundaries of the established system. Therefore the problem of a bounded semiotic does not depend on closure in itself, but it is rather strictly dependent on the way closure is established. In order to give an appraisal of the explanatory efficacy of a semiotic system, one has to wonder whether the borders of the closure are adequate or not to the object of study.

Now, the objections first embodied approaches to cognition raise against the functional-computational account of meaning-production, seem to concern the way those theories individuate the level of pertinence of meaning-acquisition in cognitive processes. By individuating the level of pertinence of meaning in the formal relations of a closed semiotic system, whose boundary is the skull, those approaches are not able to account for the acquisition of meaning in cognitive activities performed by an embodied cognitive agent. They fail to give a *ground* to the symbolic cognitive system (Anderson 2003; Harnad 1990; Glenberg, Robertson 2000; Niedenthal, Barsalou, Winkielman, Krauth-Gruber, Ric, 2005). This makes their model inadequate to its object of study: human cognition.

In particular, by considering the body as transcendent that bounded semiotic system, and defining the closure of the semiotic system *before* the cognitive agents' *body*, they fail to account for the explanatory value of mental representations "in real-world terms, for the agents that deploy them" (Robbins, Aydede 2009: 4), when they perceive the world and act upon it.

The first attempts to re-situate cognition within the context of human experience consist in moving the level of pertinence of the semiotic system so that it includes the body, which is considered to be the matrix of meaning-acquisition and cognition more in general. This implicitly entails a relocation of the body in the theoretical space of the research: it stops to be a vague and blurry materiality that the structural grating cannot grasp, and it becomes the primary source in virtue of which symbolic entities acquire their meaning.

The ultimate goal of the inclusion of the body within the closure of the semiotic system, is to find that grip on the reality of experience cognitivists' formal language-like mental system misses. Indeed, the core idea of this embodied or grounded approach to cognition is that mental representations can play a meaningful role in a cognitive system to the extent that their very constitution, their being *embodied representations*, immediately points to what cognition is: a skillful bodily activity.

Now, to understand this point -that is the way semiotic items of a mental symbolic system acquire a grip on the reality of human cognitive experience because they are thought as constitutively connected and dependent on bodily activities- it is worth to take a step back towards the broader theoretical frame in which the relation between meaning and body is the central subject of study: cognitive linguistics. Later, a short sketch of what embodied representations are and of the way those ones play a certain role in the theoretical history of embodied approaches to cognition will be provided.

I.4 Cognitive linguistics. Looking for an embodied matrix for meaning-acquisition

Cognitive Linguistics, a research field born in the US in the 80's, is grounded on the idea that cognition has to be found where there is language, or where there is a language-like structure or function.

Generally speaking, the aim of this research can be summed-up in this way:

> "Cognitive Linguistics focuses on language as an instrument for organizing, processing and conveying information. [...] Language, then, is seen as a repository of world knowledge, a structured collection of meaningful categories that help us to deal with new experiences and store information about old ones" (Geeraerts, Cuyckens 2007: 3-5)

That is, cognitive linguistics' aim is not studying language as a mere communicative means. Rather, this research aims to find a way to explain all kinds of cognitive phenomena, namely different ways in which cognition has a grip on reality, by focusing on language. This one is not considered to be independent from cognition but, on the contrary, is said to be constitutive of it (Lakoff 1990: 40).

At first glance, the core claim of cognitive linguistics seems to point to the same direction the classical cognitivist paradigm headed to, namely the idea that, to understand what the mind is, the scientist has to look at the way its functions unfold linguistically.

Nevertheless, this similarity makes sense on a very general level only. Indeed, the very birth of cognitive linguistics as a new research field is set forth by the attempt to overcome some problems the historically prominent position in linguistics -namely generative grammar, first developed by Chomsky- entailed. In particular, cognitive linguistics claims that semantics should be radically reassessed: another way to explain how linguistic signs acquire their meaning should be discovered.

Generative linguistic is grounded on a "generative commitment". This "generative commitment" consists in claiming that language should be described as a "system of combinatorial mathematics", as a set of "formal grammars" (Lakoff 1990: 43). The main aspect of language taken into account in this theoretical framework is then the syntactical one. Nevertheless, between the 60's and the 70's, many linguists of the Chomskyan tradition pushed the generative project towards "Generative Semantics" (Taylor 2007: 571).

By endorsing the core idea of Generative Grammar, generative semantics tries to explain meaning by focusing on a description of transformational rules that convert deep syntactic structures in a "surface" structure, which is the semantic one. The core idea of this formalist approach to meaning consists in the postulation of two distinct and autonomous modules, in which there are different production rules (Chomsky 1965): one for syntax (a set of strings of uninterpreted symbols) and one for semantics.

Meaning-acquisition is explained in mechanistic terms, and this recalls the computer metaphor (Lakoff 1987: 302 - 303) previously taken into account: the semantic module takes the syntax module as its input, it performs an algoritmical translation of the occurrences of this formal structure, and produces linguistic meanings as outputs of this process. Linguistic meanings are created through this process of linking two different formal systems, which translates the entities of the first domain into the second one, following appropriate rules. These two formal systems, even if they interact in a certain way, keep on being two different systems.

By considering this idea of the separation between syntax and semantics, cognitive linguistics makes its objections to this way to conceive of semantics. The generative model, even if it tries to draw a possible link between these two domains, is committed to a counterintuitive and empirically blamable assumption: the idea that syntax logically comes before semantics, namely the idea that it plays a foundational role. Syntax founds semantics, which is considered to be a superficial structure, an effect of a deep structure that causes it, in a certain sense.

To Lakoff, this idea is not motivated by the observation of linguistic phenomena, by an analysis of what humans actually do in their linguistic practices. It rather depends on the a-critical endorsement of a linguistic product as a core heuristic tool in an epistemic frame: the idea that natural language can be metaphorically modeled as a mathematical language. The core claim of this objection is that any attempt to explain semantics within a generativist framework is unsuccessful, doomed from the start, because this explanation does not understand that the justification of the reasons it provides does not depend on the consistency of reasoning itself, namely on the logical consistency of the argumentation. Justification comes rather from a space of reasons that comes before the argumentative structure the generative model appeals to. This space of reasons is that built on the mechanistic metaphor of language, which shapes scientific practices at any level.

The problem with that does not consists in the use of metaphor in itself, but it depends on the use of the wrong metaphor. In fact, by making use of the

"computer metaphor" to account for language, the generativist project underplays the most important feature of language: language is primarily and originally a *semantic realm*, because i) it is an expressive means of thought, ii) it is also the matrix, the ground from which concepts arise and constitutively depend on.

Contrary to what Chomsky claimed (Chomsky 1965), Cognitive Linguistics states that language is not separated from the faculty of conceptualization; it is not "encapsulated" in a module that is autonomous from the conceptual one, which primarily deals with meanings, but it rather lays the foundations of it.

In this peculiar theoretical framework, in which language and concepts are thought to be constitutively tied, cognitive linguistics develops the first attempt to *give a body to meanings*.

The core idea of Cognitive Linguistics, first developed in two famous books by Lakoff and Johnson, *Philosophy in the flesh* (1999) and *Metaphors we live by* (1980), is the following one. Metaphorical language, that kind of figurative language that linguistics usually studied at a rhetorical level, actively structures humans' whole cognitive life, and it does so because metaphors have an embodied ground.

Lakoff and Johnson describe metaphor as a system in virtue of which one aspect of a concept is comprehended in terms of another concept or another set of concepts: metaphor is the linguistic (and conceptual operation) by means of which two semantic fields are unified. Metaphor linguistically materializes the conflation of two semantic and conceptual domains.

I give an example in order to explain that better. In our everyday linguistic practices, it is very common to understand the concept of "arguing" by making use of the concept of "battle" or "war": argument is war. We use a variety of expressions, such as:

> "Your claims are *indefensible*.
> He *attacked every weak point* in my argument.
> His criticisms were *right on target*.
> I *demolished* his argument.
> I've never *won* an argument with him.
> You disagree? Okay, *shoot*!
> If you use that *strategy*, he'll *wipe you out*.
> He *shot down* all my arguments"
> (Lakoff, Johnson 1980: 4)

It is important to notice that we do not just talk about arguments in terms of war or battle. Our experience of arguments, what we do when we argue, is a battle. We lose or win arguments, we see the person we argue with as an opponent, we attack her claims and we defend our position, we use argumentative strategies, and so on.

This example shows what we do when we make use of a metaphorical concept: we use "war" as a metaphor to understand what we do in the (semantic) realm of "arguing" and, at the same time, we structure our linguistic practice metaphorically. We make use of linguistic expressions as conceptual tools, and we do that in a natural and quasi- immediate way because the mapping between conceptual realms metaphors produce is not arbitrary. We do not need to perform implicit inferences to subsume a linguistic occurrence under a pattern of meanings and possible linguistic uses set by convention. On the contrary, our use of these expressions is quasi-immediate, because metaphors our linguistic experience unfolds by are motivated by that thing that makes us "us", by what we cannot avoid to deal with all the time: our body.

Linguistic metaphors our conceptual system is built on are motivated by what we do by means our bodies, or, more precisely, by means of a particular kind of knowledge that blends propositional-like knowledge with a sort of practical knowledge. This hybrid knowledge includes

a) the subject's perceptual experience of her body;

b) the subject's conceptual understanding of her body, which includes folk and/or scientific knowledge of her body;

c) the subject's emotional attitude towards her body (Gallagher, Zahavi 2008: 146).

This kind of knowledge is what Lakoff and Johnson describe with concept of *image-schema*. Image-schemas are "dynamic analog representations of spatial relations and movements in space derived from perceptual and motor processes" (Gibbs, Colston 1995: 349), such as in-out, pushing-pulling, near-far, front-back. They are representations of sensorimotor and perceptual possibilities of the body, namely representations of what our bodies can do in the space.

According to Lakoff and Johnson, metaphors are built on the recurrence of such image-schemas. Their core idea is that linguistic experience, and then conceptual experience too, unfolds through the representation of bodily relations and possibilities of embodied actions.

The argumentation they offer can be unpacked in this way. Since language is grounded in experience, and since any kind of experience humans have is mediated by the bodies we have, a grip for linguistic uses and concepts have to be found where the body is or, more precisely, where the body *makes sense*. This "making sense of the body" is precisely individuated in image-schemas. Those ones, by representing aspects, features and possibilities of the body, give the agent a cognitive access to the body. In doing so, they consequently light up the *why* of the metaphors our linguistic and conceptual experience live by: they cognitively penetrate the "average-everydayness" of our bodies, which motivate all metaphorical experiences.

The core idea of this approach to cognition is that image-schemas, by representing and deploying that hybrid knowledge of the body the subject has, "explain" how the meanings human cognitive life unfolds by make sense for creatures with this kind of body.

> "Concepts of front and back [...] make sense to beings with fronts and back only. If all beings on this planet were uniform stationary spheres floating in some medium and perceiving equally in all directions, they would have no concepts of front and back" (Lakoff, Johnson 1999: 34).

Image-schemas give a representation of the way our bodies function as a *constraint* for our cognitive experience, understood in its linguistic and conceptual aspect.

To conclude this short sum-up of the core ideas of Cognitive Linguistics, it can be said that this research field matters to the theoretical development of embodied accounts of cognition because it is the first attempt to give back to language the substance the generative/cognitivist tradition stole to it. By conceiving of the body as a "semantic engine" (Gallagher 2011a for this expression), namely as something that functions as a constraint for meaning-acquisition and conceptual development, cognitive linguistics gives a living and embodied response to the problem of "meaning grounding" and "symbol grounding" standard explanations in cognitive science suffered from. The process thanks to which mental symbols acquire their meaning, namely their grip on reality, is no longer described as an obscure disembodied computational procedure, which maps the relation between two distinct formal systems, but it is rather conceived to be radically grounded in the experience of the body.

The body comes into the process of meaning-acquisition to the extent that it is represented by that particular kind of hybrid representations that image-schemas are. Metaphors make use of those hybrid representations in order to schematize and to make meaningful embodied experience, and the linguistic and conceptual ones as well. Indeed both kinds of experience are said to be metaphorically grounded.

To express the core idea of cognitive linguistics by a motto I would say:
«*The mind is where embodied metaphors, built on image-schemas, structure our cognitive processes*»

I.5 Grounded Cognition. Embodying modal representations.

The "psychological" counterpart of this project aimed at giving a body to meanings developed in linguistics is that research field known as "grounded cognition"[9] (see Barsalou 2008).

Grounded cognition begins its argumentation by pointing out that contemporary cognitive scientists, namely those who first took part in the "cognitive revolution", develop radically new approaches to mental representations, which strongly contrast with pre-twentieth century thought (Barsalou 1999: 578; Barsalou 2008: 619). Before that said cognitive revolution, the most dominant view of cognition held that higher order cognition is inherently perceptual: it relies on mental images that produce a synthesis of sensory perceptions. With the birth of cognitive sciences, which is heavily influenced by developments in logic, linguistics, statistics and computer science, the way scientists think of mental representations as cognitive media radically changes. As explained previously, the formalist trend at the core of cognitive sciences pushes scientists to think of mental representations as "sanitized" representations[10], namely as representations perceptual features are completely cleaned up from.

According to grounded cognition, those approaches to mental representations not only suffer from the problems cognitive linguistics wants to find a solution to, namely the idea that, by conceiving syntax and semantics as separated systems, those theories lead to an obscure and unsatisfactory account of meaning. Indeed, standard explanations of mental symbolic systems also claim that the mental module designated for semantic knowledge is autonomous from the brain's modal systems, such as that of perception (e.g.

---

[9] It is worth to make clear that the expression "grounded cognition" refers to a wide set of theories, which focus on different aspects of cognition, such as memory (e.g. Glenberg 1997), social cognition (e.g. Goldman 2006), the relationship between language and action (e.g. Glenberg, Kaschak 2002; Pulvermüller, Hauk, INikulin, Ilmoniemi, 2005). Nevertheless, in this paragraph I will focus on one problem only, namely the symbol grounding problem, and I will take into account Barsalou's approach to grounded cognition. In particular I will discuss the theory of perceptual symbols, which aims to account for conceptual embodiment.

[10] For the sake of accuracy it is worth to notice that the expression "sanitized representation" was coined by Gallagher (Gallagher 2011a) to refer to Goldman and de Vignemont's B-formats (body representations). The way I use this phrase in this paragraph has no reference to the objections Gallagher makes to this approach to embodiment. The use of this expression has rather a pure heuristic value in the economy of my exam of Barsalou's version of grounded cognition.

vision and audition) action (e.g. movement and proprioception) and introspection (e.g., mental states, affect; Fodor 1975; Fodor 1983).

Standard approaches to cognition endorse *a-modal* theories of knowledge: cognition is said to depend on principles that are completely different from those which rule perception. Perception, exactly because it rests on mechanisms different from those higher order cognition unfolds by, is said not being part of the realm of "fully-fledged cognition", that is knowledge that takes propositions as models in order to understand cognitive processes.

To put it in other words, according to standard approaches to the mind, perception is something different from "real cognition" because they postulate a priori that mental symbols, namely what ontologically constitutes the mind, have a non-perceptual nature. They are rather defined as a-modal, namely as symbolic-systems that bear an important relation to words and language. To explain this point Barsalou writes:

> "Just as language processing is assumed to involve the sequential processing of words in a sentence, so conceptual processing is assumed to involve the sequential processing of amodal symbols in list-like or sentence-like structures" (Barsalou 1999: 579).

A-modal symbols, in the same way words of sentences are not related to their referents by systematical similarities, do not stand for the perceptual states that take place during cognitive processes according to an analogical relation, which is what, in semiotic terms, is an iconic relation that maps the formal similarities between two items (CP 2.778). Those theories, by modeling any mental entity on a-modal, language-like symbols, not only do not account for the "material", embodied ground of mental representations, but are also unable to give a successful explanation of the way mental symbols can grasp perceptual reality. By focusing on abstract and formal properties of mental symbols only, those accounts seem unable to explain how those formal structures are mapped back into perceptual states and entities of the world. That is to say: by modeling those symbols on words, which entertain a mere arbitrary connection with the objects they stand for, the theory holds that the relation mental symbols entertain with perceptual states is just arbitrary.

According to supporters of grounded cognition, by conceiving the relation between mental symbols and sensorimotor states in that way, those theories do not actually account for cognition as it is experienced by the cognitive agent when she takes part in the cognitive process, sensing and perceiving the world.

Then, the core aim of grounded cognition is to find a different way to account for mental representations. To some extent, this approach to cognition seems to go back to the pre-twentieth century idea that knowledge (included higher-order cognitive skills, such as language and conceptualization) widely makes use of mental images, that is *picture-like representations*.

The main idea that this new explanation pushes forward is the following one. The concept of mental representation is useful to explain cognition if and only if it accounts for agents' cognitive experience. Since cognitive experience primarily unfolds in a perceptual way - that is, our primary encounter with the world takes place by means of the sensory apparatus- the symbols the theory should rely on should have a *perceptual nature* and they should be *modality specific*, namely dependent on the sensory modality information comes from.

Those ideas merge in what is known as the theory of *Perceptual Symbols Systems - PSS*.

To begin my discussion about PSS, now I consider the core claim of this new theory about the format of mental representation by contrasting it again with standard approaches to cognition.

Standard approaches to cognition account for cognitive processes that "begin" with sensory inputs and "end" with the conceptual organization of information in the following way (Barsalou, Simmons, Barbey, Wilson 2003: 85).

i) The cognitive agent's sensory apparatus is stimulated by a physical stimulus.
ii) Information about the physical stimulus travels up sensory channels, in this case the visual one.
iii) Neurons aimed to feature-mapping fire and produce a sensory representation of the stimulus.
iv) This modal representation of perceptual states is transduced into a non-representational perceptual format. Non-perceptual formats can be: a) a representation in which some features of the physical object are listed (feature list); b) a semantic network, which represents semantic relations between concepts using nodes and vectors (it is something like a mental graph); c) a frame, which is a set of informational slots, whose values and types are specified.

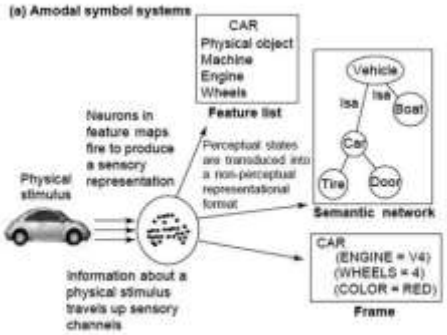


Figure 1.
[Picture taken from Barsalou, L., Simmons, W., K, Barbey, A., K., Wilson, C. D. (2003)]

This is to say that according to those standard accounts of mental representation, representations belonging from modal systems are *transduced* into a-modal symbols that represent knowledge about experience in semantic memory (Barsalou 2008: 618). Representations that stand for their objects accounting for the variables each sensory modality (e.g. visual, haptic, auditory) conveys are *re-described* or *translated* in an a-modal way, abstracting from those variable aspects. The final product of this operation of transduction is an invariant structure that represents a category of objects (a concept) in all contexts. It is a *non-context sensitive* representation the cognitive system makes use of and re-uses in its overall dynamics, never relying on memories of the sensorimotor states that transduction redescribes. It is a *disembodied representation*, since it does not bear any relation with the sensorimotor state that caused it. It is a *static* representation: it is a-temporal because it is defined without any reference to real-time cognitive activities; it is an invariant, *discrete* symbol (Barsalou 1999: 584).[11]

According to supporters of grounded cognition, the problem with this kind of account of mental representation concerns the effects produced by the operation of transduction, namely what I would call "sanitizing process of mental images". Standard views of mental representation postulate that the overall dynamic of cognitive processes has to be described in terms of a-modal representations. Those ones are symbols the cognitive system makes use of when the stimuli that caused them (a sensory state) is absent. Any kind of cognitive activity is said to unfold by recalling those redescriptions of sensorimotor states. Those last ones, taken in themselves, namely as items not

[11] For a more detailed description of features ascribed to a-modal representations see Barsalou, Prinz 2000.

processed by the a-modal translation, are said to play no relevant role in cognitive dynamics such as language, memory and thought.

The problem with this account of cognitive processes consists in the fact that the explanation, by stating the primacy of this kind of representations in cognitive processes (a-modal symbols are said to rule any aspect of human mental life), seems to be unable to account for perceptual cognition in a changing environment. How invariant, non-context sensitive and non-modality specific representations can explain the way the perceiver is able to cope with an environment that constantly changes? The features standard views ascribe to representations do not seem to fit with behavioral evidence: cognitive subjects are able to cope with their environment, by attuning their behaviors to the situations they face. Invariant and non-context sensitive representations seem unable to account for the adaptivity of the agent's behavior, then it is not clear why evolution should have selected mental representations as a "central design function for cognition" (Barsalou, Prinz 2000: 58). "Sanitized representations" do not have adaptive function: they do not guide action.

In contrast with this a-modal approach to cognition, PSS posit a kind of representational process whose core features can be explained in the following way.

i) As in standard approaches to mental representations, at the beginning neural representations represent objects in vision.
ii) Nevertheless, contrary to old-fashioned approaches to mental representations, grounded cognition claims that visual representations, instead of being transduced into a-modal descriptions, are captured by neurons in nearby association areas.
iii) Later, when the sensory input is no longer present, the system is not said to recall those "sanitized representations" standard approaches postulated. Rather, the areas nearby the visual area *re-enact* visual states.
iv) This re-enactment of perceptual states contributes to the overall dynamics of cognitive process. For example, there is re-enactment in memory, language and thought.

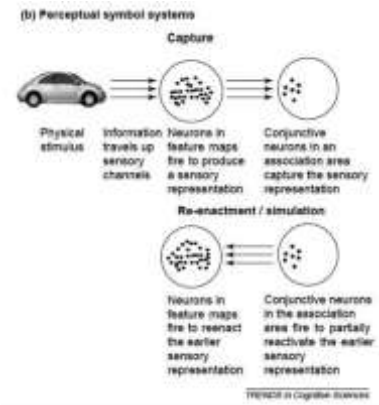


Figure 2.
[Picture taken from Barsalou, L., Simmons, W., K, Barbey, A., K., Wilson, C. D. (2003)]

As this brief sum-up shows, the main difference between the two approaches taken into account consists in a theoretical shift: transduction is substituted by the concept of re-enactment. Broadly speaking, the word "transduction" is meant to refer to an operation that produces a passage of an X in a given field to another field, where the two fields or domains are heterogeneous. The outcome of transduction is an irreversible change of the nature of X. For example, think about the membranes of microphones. When an analogical transduction takes place (namely when an X of a given field is mapped into another field preserving the internal relations of X in the first domain into the second one) the mechanical energy produced by the speaker's voice is transformed into electric energy by the microphone membrane. The microphone membrane works as a transductor of energy (Paolucci 2011: 408): it turns mechanical energy into electric energy. These two kinds of energy are different Xs, but their structures preserve an analogical relation: mechanical energy in the first domain is what electric energy is in the second domain.

At first glance, it seems that, according to Barsalou, the problem with transduction concerns the domains the transposition deals with. The outcome domain in the case of the transduction of sensory signals, as it is conceived by a-modal approaches to representations, is the crystalized domain of abstractness and a-modality. The operation of transduction, in the cases critically taken into account by Barsalou, is cleaned up by its analogical features. Indeed, when sensory signals are transposed at the symbolic level, they seem preserving no significant relation with what they were before transduction. They do not bring traces of what caused them, namely the perceptual properties of the objects the agent deals with while she performs her actions. Those features are dynamical features, not only because they have dynamical effects of the agent's bodies (for example, think of the microkinesis of the body) but also

because they are part of a perceptual dynamics in which the agent does not perceive neutral objects, but she perceives objects she can do something with in that context (e.g. the example of the chair in Barsalou, Prinz 2000).

Senses mean something in cognitive processes because they afford action. In the case of symbolic representations the issue of meaning is not considered in this way, it is not conceptually linked to the problem of action. Indeed, symbolic representations are tokens of conceptual types. They are part of a representational relation because they are occurrences of their types: they acquire meaning because they can be subsumed into a category. Therefore, the kind of relation that symbolic tokens have with their types is a *passive* relation: they are objects of a subsumption, and not agents of a relation.

Barsalou's PSS are aimed at justifying and explaining the analogical relation that connects what the agent senses (e.g. possibilities for action) in her perceptual experience and what happens inside of her head. Moreover, Barsalou's concept of perceptual symbol is aims to explain how the sensory domain and the conceptual one can be considered to be part of the *same system*.

The idea of re-enactment is central in this project. The re-enactment activity Barsalou talks about in order to account for representational activity in the absence of the stimuli that first caused the representation-token is a specific kind of *simulation*. It is the partial re-enactment of the *active features* of sensorimotor representations performed by a neural activation (Barsalou, Simmons, Barbey, Wilson 2003: 85)[12] when the stimulus is absent. This simulation of perceptual states performed by the neural area nearby the one activated during the very perceptual and motor state of the cognitive process is what the expression "perceptual symbol" refers to.

Perceptual symbols represent the object they stand for, that is the perceptual state, by *re-activating* the perceptual/motor state that caused the first representation in nearby association areas. They "channel" action and perception into a neural area of the same system in which sensory signals occurred, and they preserve the perceptual and active features of the stimuli.[13] Nevertheless, even if re-enactment entails a passage from a neural area to another one, this passage has a different nature from the mechanisms described

[12] See also Barsalou (2009),"Simulation, situated conceptualization, and prediction", *Philosophical Transactions of the Royal Society*, 364, pp. 1281-1289 and Barsalou (2008), "Grounded Cognition", *Annual Review of Psychology*, 59, pp. 617-645 for a detailed description of empirical evidence of neural re-enactment.

[13] This operation is known as "neural reuse": neural circuits established for one purpose acquire new uses after an initial or original function is established (Barsalou 2016; Anderson 2010).

by standard views of representations: it does not entail transduction because perceptions and simulations of perceptions are part of the same system. Moreover, the operation of simulation Barsalou talks about has the following features.

First, it is *active*: re-enacting means *re-acting out*, that is representing something through of action (as it happens in theatre performances, for instance).

Second, this simulation is *dynamic* and *not discrete*. Perceptual symbols are not something like feature lists, frames or semantic networks, namely static and invariant representations of a set of objects (concepts). They are rather conceived of as associative patterns of neurons whose activation may vary widely (Barsalou 1999: 584) according to the context.

They are *context-sensitive* and *flexible*: they are like "tracking devices" (Barsalou, Prinz 2000: 63), which track features of the object they stand for by representing them according to what the context requires.

They are *modality-specific*: they are representations that stand for their objects by entraining an *analogical relation* with the perceptual activation that caused them. Neural areas activated when the stimuli is absent and present are not exactly the same; nevertheless they are responsible of the same mechanisms. Perceptual symbols are represented in the *same system* of the sensorimotor/perceptual state that caused them, that is why they are not said to be transduced: there is not a passage from a system to another (this is what the word "transduction" suggests), but there is a re-enactment of sensory processes that occurred in the same system, namely the perceptual one. To explain this point Barsalou writes:

> "The neural systems that represent color in perception, for example, also represent the colors of objects in perceptual symbols, at least to a significant extent. On this view, a common representational system underlies perception and cognition, not independent systems. Because perceptual symbols are modal, they are also analogical. The structure of a perceptual symbol corresponds, at least somewhat, to the perceptual state that produced it". (Barsalou 1999: 578)

This is to say that perceptual symbols, namely mental representations the cognitive system uses also when the cognitive process does not deal with perceptual or sensorimotor tasks, are embedded in the same neural system responsible for perceptual states. This suggests that representational

mechanisms that guide perception and higher order cognition have the same structure.

To offer some philosophical considerations about this description of mental representational activity, it can be said that because concepts and sensorimotor/perceptual states are said to be embedded in the same system, PSS theory is able to guarantee the explanation an embodied and active *ground*. By describing mental symbols in this way, and by substituting the concept of transduction with that of simulation or re-enactment, PPS not only explain the way representations are embedded in neuronal areas, but it also explains how representations relate to the embodied activity of the perceiver. The brain, by representing dynamic processes analogically, "re-performs internally" what the cognitive agent does in her cognitive experience. At the same time, it displays cognition, perception and action as part of the same neural system.

In this way, the explanation of mental activities offered by grounded cognition seems to take the issues cognitive linguistics deals with seriously. As cognitive linguistics does, grounded cognition in Barsalou and colleagues' version of embodied cognition tries to found an embodied matrix for meanings and concepts. Moreover, it gives an embodied account of the continuity of cognitive skills (e.g. reasoning, thinking, perceiving, memory), and it does so by giving an epistemic primacy to the cognitive agent's body.

To conclude this paragraph, I would say that, if a motto for this approach to cognition should be coined, it would sound like this: *«The mind is where there is a sensorimotor activity or a re-enactment of it in terms of perceptual symbols».*

I.6 The body's *βίος*: towards the very materiality of the body.

In the last paragraphs, I took into account two interpretations of embodied cognition.

What those approaches share is the philosophical problem they come from: the symbol grounding problem.

This motivates the way they account for embodiment in mental processes. The mottos I wrote in order to sum-up the core claims of those views of cognition were:

i) «*The mind is where embodied metaphors, built on image-schemas, structure our cognitive processes*».

ii) *«The mind is where there is a sensorimotor activity or a re-enactment of it in terms of perceptual symbols».*

As the two sentences show, the idea of embodiment that Cognitive Linguistics (in Lakoff and Johnson's version) and Grounded Cognition (as endorsed by Barsalou) give is strictly connected to semantic and conceptual issues. Cognition (and consequently the mind) is said to be embodied because meanings and concepts take the body as their "matrix of emergence".

Even if in different ways, these two approaches still make use of the concept of representation as an explanatory tool to have an access to the body. Obviously, the idea of perceptual symbol and that of image-schema are peculiar ways to think of cognitive representations, which differ from standard views of mental representations, namely from representations modeled as "words". In particular, perceptual symbols and image-schemas are attempts to account for cognition as active and embodied. As already shown, perceptual symbols are neuronal patterns that re-enact sensorimotor states in the absence of the sensory stimuli, by preserving an analogical relation with the sensorimotor state that caused the activation of the neuronal areas nearby. Image-schemas are what I defined "hybrid representations" (they are also called weak representations[14]). They are not proper sensorimotor states (Gibbs, Colston 1995: 349), but analogical representations of them. Moreover, this *perceptual* (and emotional) *implicit* knowledge about the body is combined with the subject's *conceptual* understanding of her body. This kind of representation, as in the case of perceptual symbols, acts as a mediator in the process which structures higher order cognitive processes (i.e. language and memory).

[14] See Gallagher 2011a.

Cognitive linguistics' approach (in Lakoff and Johnson version) to image-schemas particularly stresses the conceptual aspect of them. They are through to be conceptual (Lakoff, Johnson 1999: 34) and what is interesting is that when Lakoff and Johnson discuss the inseparability of categories, concepts and experiences, they define concepts as "neural structures that allow us to mentally characterize our categories and reason about them" (Lakoff, Johnson 1999: 19). Therefore, even if in other publications they state that the only concept of representation they endorse is not an "inner mental entity" but a "flexible pattern of organism-environment interaction" (Lakoff, Johnson 2002: 249-250, as cited in Gallagher 2011a: 64), Lakoff and Johnson seem to still refer to a sort of embodiment "neurally mediated". If image-schemas are enabling conditions for the metaphorical thought, and if they are thought to be conceptual within an explanatory frame in which concepts are (sometimes) defined as neural structures, then it is possible to claim that, somehow, this approach to the body can be considered as a *weak* version of embodiment (Gallagher 2011a).

Now, some embodied approaches to cognition that come from phenomenology (Gallagher 2005), philosophy of psychology and biology (Shapiro 2004), and biology (Chiel, Beer 1997) reproach to approaches to cognition that look for the cognitive role of the body in embodied representations to underplay the role of the body "before and after the brain". That is to say: those theories are said to underestimate the role of the body (conceived in its anatomical, motor, and chemical aspects) in pre-processing and post-processing information in the cognitive system (Gallagher 2011a).

According to those approaches -which do not constitute a philosophical school in a proper sense, but that nevertheless share the idea of "biological embodiment"- in order to actually account for the role of the body in cognitive process, embodied theorists should consider its *very materiality* as a factor that does *the most part* of the work in cognitive processes.

Reassessing what Gallagher points out (Gallagher 2011a: 61) quoting Shapiro (Shapiro 2004: 190), it can be said that the idea of "biological embodiment" holds that the materiality of the body not only constrains cognitive processes (i.e. the body affords some kinds of cognitive operations and prevents the cognitive agent from performing other kinds of cognitive activities). The body, precisely in virtue of its *extra-neural* "material constitution" (anatomy), its "dynamical constitution" (motor aspect), and its "chemical constitution" (see for instance hormonal processes), should be taken to be a constitutive part of the cognitive system. This is to say, those factors should not be considered as something that has just a *causal influence* on the

mind (also cognitivism recognized that –Shapiro 2011: 159); rather, they are *constitutive* of the mind. Indeed, changes and differences in anatomical, chemical and motor factors produce changes and differences in the *effects* of the cognitive activity (i.e. cognitive contents) and also in the *modality* in which cognition unfolds (i.e. fluency in performing the cognitive task). That is why they should be considered to be constitutive of the mind.

In Gallagher's article previously quoted, empirical evidence to support this philosophical claim is given. For instance, he cites research done by Roll and Roll (Roll and Roll 1988: 162) that shows that changes in body postures caused by vibration-induced proprioceptive patterns change the way the environment is perceived, and how hormonal changes, visceral and musculoskeletal processes affect perception, memory and decision-making (Damasio 1994; Bechara et al. 1997).

Other empirical evidence to proof how the body plays a constitutive role in *determining* the mind's contents and the qualitative aspects of cognitive processes (modality) before neural processing can be also found in some research developed by the neuroscientist and cognitive scientist Daniel Casasanto at the psychology laboratory of Chicago University. Those experiments are meant to verify the "body-specificity hypothesis" (Casasanto 2009), or "bodily relativity". According to the "bodily relativity hypothesis",

> "to the extent that the content of the mind depends on our interactions with our environment, people with different kinds of bodies – who interact with the environment in systematically different ways – should tend to form correspondingly different neural and cognitive representations" (Casasanto 2011: 108)

The hypothesis that should be verified consists in saying that cognitive representations (namely mental contents) constitutively vary according to the cognitive agent's body. Moreover, the hypothesis entails that the way the "anatomical and motor body" affects mental contents follows a *continuous* path: this does not only shapes perceptual contents, but it also constitutively determines the contents of higher order cognitive processes.

The basic idea of this hypothesis is that anatomical and motor differences of the body do not merely correlate with cognitive differences; they rather entertain a stronger relation with mental contents: mental contents, cognitive representations, neural representations entertain a relation of constitutive dependence with bodily features.

That is why, shifting the argumentation to a more philosophical level, those ones should be considered to be constitutive of the cognitive process, and then of the mind. The body matters for cognition not only because it is represented in a certain way (i.e. PSS and image-schemas) by the mind, and then those representations deeply influence the overall dynamics of cognitive processes, but also and especially because it entertains a "more primitive" relation with cognition. Representations that are said to give a cognitive access to the body are strictly dependent on the very materiality of the body they represent. Moreover, the body seem to be considered to be cognitive also *before* it is processed by internal representations.

To understand this point, it is worth to have a look to the experimental part of Casasanto's research. In the article "Different Bodies, different Minds: The Body Specificity of Language and Thought" (Casasanto 2011), Casasanto describes an experiment made with Evangelia Chrysikou in which they studied how people think about "god" and "bad" and make judgments about those values after the ordinary function of their dominant hands has been handicapped. The experiment tested space-valence mappings (namely the association of positive or negative values and objects situated in different parts of the spatial context of the cognitive task) in a task in which healthy university students were asked to perform a motor-fluency task while wearing a clumbersome glove on their left hand or on their right one. This temporarily turned right-handed people in left-handed people and left-handed people in right-handed people. After 12 minutes of lopsided sensorimotor experience, students removed the glove. Then the cognitive task was performed again without the glove-constraint. Casasanto and Chrysikou discovered that students that had worn the left glove during the first task, still thought that "right" was good (that is, they associated "good" with "right" in their judgments about objects in the space). On the contrary, participants who had worn the right glove associated "left" with "good", as naturally left-handed people usually do.

The conclusion the two scientists drew from that experiment (whose results were compared with those from experiments in which space-valence mappings were tested in relation to stroke patients with hemiparesis on their left or right side)[15] is that higher order and complex forms of cognition (i.e. emotional judgments) are strictly connected to the very material constitution of the body (and to the way this has a great influence on motor habits). This means that the body matters for cognition, it "shapes the mind", also before complex

[15] In order to check the explanatory efficacy of the results of this experiment, see also Casasanto 2009; Casasanto, Chrysikou 2011; De la Vega, Dudschig, De Filippis, Lachmair, Kaup 2011.

representational processes take place (this can be tested by looking at the timing of emotional judgments formation). The body does not only give a ground to mental representations, but it works as a "content shaper" also before brain processing.

This emphasis on the very materiality of the body is what another theory in the "4E cognition" debate, Enactivism, focuses on, integrating it with some philosophical assumptions taken from phenomenology.

I.7 Making sense of inter-action: Enactivist Embodiment.

Until this moment I focused my attention on some ideas about cognition some approaches in the embodied cognition framework suggest. I started my discussion taking into account the "symbol grounding" problem the cognitivists' explanations suffer from (namely what lays the theoretical foundations of new approaches to cognitive science), and I summed-up the way some theories in the embodied cognition framework deal with this issue. This discussion showed the necessity to *re-embody the mind*, in order to successfully explain the *experience* of cognitive processes. Until this point of my discussion, the inclusion of just one of the said "dirty factors" (see §I.1.2) in the explanation of cognitive processes was taken into account. What can be said is that the considerations made point to a redefinition of the mind's borders by including the body (considered as a brain and as a whole bodily structure) as one of its constituents. *The mind extends into the body.*

Nevertheless, as sketched out in §I.1.2, the broader project of "4 Es" approaches to cognitive sciences points to a picture of the mind that is even broader, whose borders encompass the context where cognitive practices take place as well.

Considering the temporal development of "4 Es Cognition" it can be said that, strikingly, the inclusion of factors that are *beyond the body* within the realm of the mental is first suggested by a theory that is "fully embodied": the enactivist approach to cognition.

The roots of the Enactive Theory of cognition should be found in works by Varela, Thompson, and Rosch (*The Embodied Mind*, 1991), and by Maturana and Varela (*The Biological Roots of Human Understanding,* 1987).

As said at the end of the previous paragraph, this approach to cognition and the mind begins its philosophical considerations by claiming that cognition is fundamentally dependent on biological factors; or better, it states that the biology of a cognitive system is an unavoidable constituent of cognition.

Nevertheless, rather than focusing on the singularity of bodily factors (this is what I tried to do in the previous paragraph), the enactive approach to cognition starts its argumentation by inquiring the broader organization of biological *systems*, which are investigated by means of the concepts of *autonomy*, *sense-making*, and *enaction*.

Autonomy is the capacity of a system to specify its own laws (Maturana, Varela 1987: 48), namely to display *regular* mechanisms by means of which it shows its organization, - where the term "organization" refers to the relations

existing among the components of the system (Maturana, Varela 1987: 47). Those regularities are laws according to which processes that are part the system generate and sustain the biological system as a unity.

Despite at first glance it seems to be the case that this claim entails a conception of the body that accounts for it as a thing, as a *unitary autonomous res*, the philosophical direction enactive cognition takes is the opposite one. The claim that living systems should be considered to be autonomous does not entail a metaphysical view according to which living organisms should be conceived as autonomous *entities*. Indeed, autonomy describes the way the *processes* the organism unfolds by display recognizable features.

Varela says:

> "(...) autonomous systems are *organizationally closed*. That is, their organization is characterized by processes such that (1) the processes are related as a network, so that they *recursively* depend on each other in the generation and realization of the processes themselves, and (2) they constitute the system as a unity recognizable in the space (domain) in which the process exists" (Varela 1979: 55; italics added)

Biological systems, that is embodied systems, are said to display a coherent behavior because their organization is characterized by an *operational closure*. The system sustains itself (and consequently it is able to stand out as a unity in space and time) to the extent that processes that take place within the boundaries of the organism recursively depend the one on the others. This is to say that the system develops through the relation between interdependent processes, through the reciprocal attunement of processes I would say, in order to maintain a high degree of stability (homeostasis).

In the enactive approach to cognition, the body is precisely defined by this operational closure. The "closed" relations between its structure, namely the components and the relations that constitute a particular unity of the organization of the living organism (Maturana, Varela 1987: 47), are what allows the body to individuate itself as distinct from its immediate surroundings.

Then the body is described as a *self-individuating* (Di Paolo, Thompson 2014: 68) set of processes (not as an entity) that reciprocally define themselves. It is a body understood biologically, but this biological identification is not given by considering discrete parts (the chemical one, the motor one, the anatomical one) that constitute the biology of the system. The identification of the biological system is dependent on the individuation of the relations between the activities that constitute it, and on the appraisal of their capacity to sustain

themselves, allowing the body to show itself (namely being recognizable) as a unity of interdependent processes. The unity of the body should be individuated in the *consistency of its activities*.

It is precisely this capacity of self-individuation that makes the body *cognitive*. As some scholars claimed (Di Paolo, Thompson 2014; Froese, Stewart 2015), by individuating "the mark of the living" with autonomy, enactivism also individuates what is known in the broad literature about "4 E Cognition" with the phrase "mark of the cognitive".[16]

As Barandiaran points out (Barandiaran 2016: 2), autonomy constitutes the fundamental logic of the living: only a system that can maintain itself stable through changes in time, by attuning itself to changes, can survive. That is why autonomy lays the foundations of the *embodied feature* of processes. The body can be experienced in cognitive processes because there is life: it can constitute a lived and experiential structure because its organization guarantees the *continuity* of the living experiences cognition unfolds by (see also Varela, Thompson, Rosch 1991: 16).

This self- regulation of the system, its operational closure, also endows the organism with the ability to adapt itself to conditions that can produce changes in the system. This is to say that *adaptivity* is the counterpart of autonomy. Contextual conditions can improve or deteriorate the system, and this one, in order to keep on being a living system (a system whose functions persist and are coherent) has to find adaptive paths to attune itself to those unavoidable conditions.

Explained in this way, it seems to be the case that the environmental context plays the function of an external constraint for living systems: the environment the biological system is situated in compels the system to attune its processes to its conditions in order to survive. Nevertheless, despite this explanation that defines the context as an external, unavoidable constraint for living systems can seem intuitive, conceiving of the environmental context in which biological systems live just as a "perturbation factor" leads to a substantial misinterpretation of the enactive approach to cognition. Indeed, if it is true that in its birth enactivism defined its conceptual framework by making use of the notions of operational closure and autonomy, the meaning those concepts

[16] Even if in a different way (indeed the issue of the "mark of the cognitive" has concerned more the debate about the Extended Mind), this point is expressed in the *life-mind continuity thesis*. According to the life-mind continuity thesis "life and mind share a common pattern or organization, and the organizational properties characteristic of mind are an enriched version of those fundamental to life. Mind is life-like, and life is mind-like" (Thompson 2004: 385).

acquire in the overall enactive picture of cognition leads to think that the closure of the system does not entail that the environmental context is just a *constraint* for the cognitive system. The relation between the living system and the environment should not be conceived in dualistic terms: there is not a real internality of the embodied system and an externality of the environmental context. In fact, one of the core claims of the enactivist interpretation of cognition is the idea that cognition is an *inter-active* process defined as *enactive sense-making.*[17]

On the one hand, the *active attunement* of processes within the biological system defines the system as something that individuates itself as a unity that stands out from a background. On the other hand, the concept of enactive sense-making not only explains how the living system shows itself as cognitive, displaying its ability of making sense of changes over time, but it also justifies a Gestaltic-like conception of the relation between the organism and the background it stands out from, that is its context.

Before explaining what enactive sense-making is in more detail, it is worth to consider the relationship between figure and background in the Gestaltic perspective, in order to understand better why I claim that by means of the concept of sense-making, enactivism gives a "Gestaltic" account of the relation between organism and environment.

Gestalt theory is an approach to perceptual organization that focuses on the concept of organized structure or *organism* (Kanizsa 1988: 17). Instead of considering perceptual processes as developing in the association of atomistic sensations caused by discrete stimuli, Gestalt theory defines perception as a process of pattern recognition that takes place in time. This is to say, according to this approach to perception, we do not perceive individual stimuli, but perception is the recognition of recurrent figures that show themselves as perceptual wholes.

Considering the issue I am interested in, that is the relation between figure and background, it can be said that Gestalt theory accounts for this relation by

[17] It is worth to notice that, even if the concept of sense-making is already developed in early works that endorse an enactive approach to cognition (see among others Varela 1984), it seems to be the case that its centrality in the philosophical debate is more recent, and it is due to attempts to socially extend the enactivist perspective (e.g. De Jaegher, Di Paolo 2007) and to investigate the normative aspect of cognitive processes in depth (e.g. Barandiaran, Egbert 2014; Barandiaran 2016). Nevertheless, since this paragraph aims to provide some general insights about enactivism, for explicative purposes, here I integrate different perspectives on sense-making.

considering the idea of structure: figure and ground stand out as two individuated perceptual wholes (namely, they can be perceptually segregated) in virtue of their inter-definition. The figure stands out *as* a figure because its features contrast with those of the ground. Usually the figure is smaller than what is perceived as a background, it has a closed structure (and then it is perceived as more solid), it has a simpler structure, and so on. All the qualities attributed to the figure depend on those of the ground: they cannot be defined without a term of comparison (the figure is smaller, more solid, its structure is more closed, and so on), and then the figure cannot stand out as a figure, it cannot actually be perceived as a figure, without the ground. The same can be said about the ground. The two stand out as distinct perceptual wholes because there is something they stand out from. Nevertheless, the perceiver cannot perceive what each perceptual entity stand out from and each individual perceptual entity at the same time. The two perceptual paths are exclusive, and this is what makes the two perceptual entities individual. Nonetheless, this individuation is constitutively dependent on the existence of the "opposite" or "negative" of the figure the perceiver's attention is directed towards. The two perceptual poles theoretically constitute a unity: they constitute something coherent in virtue of an intrinsic relational structure. In this sense, it can be said that the ground is not something that is *actually* external to the figure, and the figure is not something that is actually contained in the ground. Not only it is not possible to perceive the "internal" and the "external" at the same time, but also, at the theoretical level, there is not an "internality" that foreruns an "externality" or an "externality" that foreruns an "internality". Perceptual organizations define their features and their individuality in virtue of the relation that connects them: their existence constitutively depends on this interaction.

Shifting the discussion towards the enactive conception of the relation between the living system and its ground, it can be said that, similarly to what happens in Gestalt psychology, according to enactivism it cannot be claimed that there is something like a locating structure (the context) that foreruns a located "thing" (the living system). Rather, the two inter-define themselves in their *dynamic interactions*. This is explained by means of the concept of sense-making. This concept can be unpacked in the following way:

> "sense-making is behavior or conduct in relation to environmental significance and valence, which the organism itself enacts or brings forth on the basis of its autonomy. [...] An autonomous system produces and

> sustains its own identity [...] and thereby establishes a perspective from which interactions with the world acquire a *normative status*" (Thompson, Stapleton 2009: 25, italics added).

Sense-making is a relational concept coined in order to account for the way the organism, in virtue of its actions, behavior or conduct, brings forth environmental conditions suitable to maintain its self-persistence or integrity (autonomy). The basic idea of sense-making is that the organism, by means of its actions within a context, creates a salient space -a space that *makes sense*- in its world (*Umwelt*). It *enacts* its own Umwelt.

"Enacting" means "acting something out" (Noë 2004: 1), "acting from within", "establishing a law" (Barandiaran 2016: 2), "bringing forth" (Varela, Thompson, Rosch 1991: 149): the organism, in virtue of its behavior, defines a salient environment that stands out from "the rest of the world" because it has a significance, a meaning for the cognitive system. This significance precisely rests in the capacity of the Umwelt to solicit a living system's actions, which are aimed at maintaining the unity of the system. Then it can be said that on the one hand, the living system defines its Umwelt by acting it out; on the other hand, the Umwelt defines the living system by providing the meanings according to which the organism's activity unfold. The context where the embodied system lives ceases to be the theatre of the agent's actions and becomes a space of valences and saliences the agent actively creates dealing with what the world offers.

To explain the peculiar relation between organism and environment more clearly, it can be said that on the one hand, the Umwelt stands out as something different from the cognitive agent, to the extent that it displays problematic situations in which the system shows up as an agent (namely it shows its agency, its capacity to act), because it acts upon something in order to bring forth conditions suitable for the stability of the system. On the other hand, the Umwelt shows itself as a "familiar space", an "affordable space" in which the agent can display intelligent behavior because the Umwelt, by displaying the meanings the agent needs, is perceived as always and already meaningful and cognitively manipulable. This familiarity makes the living organism perceiving the Umwelt as *continuous* with its actions.

That is why it cannot be said that the Umwelt just acts as a "locating" factor in relation to the living system. The agent is not just located in an environment, but it acts it out: the environment conceived as an Umwelt does not exist before the actions of the system. On its turn, the Umwelt makes the living system existing as a cognitive system: it displays the meanings, the valences and the

values the system deals with in its activities of self-individuation and self-sustainment.[18]

At this point, it would probably be clearer why I claimed that autonomy (conceived as the integrity of the living system) provides the "mark of the cognitive" in the enactive approach to cognition. The concept of autonomy entertains a relation of interdependence with those of sense-making and enaction. Indeed, those last ones are understood as repeated patterns of actions (which unfold *regularly* and *habitually* in time - Barandiaran, Di Paolo 2014: 6) that bring forth conditions able to guarantee the continuity of life, which, as explained before, is interpreted as co-extensive to cognition.

Sense-making and enaction are considered to be "cognitive operations" because they endow the world with a *normative* status, and normativity is what usually allows philosophers to define something as cognitive. Indeed, they "create" an Umwelt, a system of meanings and "life-values" according which actions can be *judged* as adaptive or maladaptive, good or bad for the preservation of life-mind in that subjective world.

As Rowlands notices (Rowlands 2012), broadly speaking it can be said that within the frame of philosophy of knowledge, it is commonly endorsed that something can be considered to be cognitive if it entails a normative claim on the world. This is to say that cognition "says" how the world ought to be: cognition has to display what is *supposed to* produce it.

To explain this point I consider an example taken from an article by Mark Rowlands, which does not precisely endorse an enactive perspective as the one I am currently taking into account, but that can be still useful to make my point clear.

Rowlands says:

> "Not everything that actually *does* produce my belief that there is a horse in front of me *should* produce this belief: the donkey that is in front of me and is, in fact, causally producing my belief should not do so. It does produce the belief, but it shouldn't" (Rowlands 2012: 134)

[18] It is worth to notice that both classical enactivists (e.g. Varela, Maturana, Thompson and Rosch) and contemporary enactivists (e.g. Noë, Barandiaran, Di Paolo, Gallagher, Rietveld) explain this point by recalling Merleau-Ponty's phenomenology. According to that perspective, the relation between the world and the agent is described in terms of inseparability (the subject is conceived as a project of the world, and the world as a projection of the subject –Merleau-Ponty 1962: 439) and of transitive and circular action (the organism is related to its *milieu* because it affects it through its actions and the *milieu* affects the organism because it "responds" to the organism's general attitude towards it – Merleau-Ponty 1963: 148).

Cognition has to be found where there is a normative domain that allows us to judge if a mental state, behavior, or action[19] should have been produced or not. This is what distinguishes cognition from a mere *fact*. Facts just happen, they cannot be judged as true or false, they cannot be disputed or rejected because they have just the modality of existence. They are not produced according to a law that displays them as connected, gathered within the boundaries of a certain domain endowed with an internal consistency. In contrast, something can be defined as cognitive because we have *reasons* to judge it as true or false, successful or unsuccessful, useful or not useful, and we can do that because there is a normative domain that makes cognitive judgments possible, by defining the lawful relations among the "entities" of this normative domain.

Discussing the enactive approach, it seems to be the case that this normative dimension cognition rests on is co-extensive with biological normativity (Barandarian 2016: 7). This naturalistic account of normativity is dependent on the endorsement of the "life-mind continuity thesis" that, as previously pointed out, lays the foundations of the enactivist perspective. Actions can be considered to be cognitive because they can bring forth certain conditions in the world that can be *judged* as "good or bad, appropriate or inappropriate", adaptive or maladaptive, as viable or non-viable ways to maintain the biological integrity of the system (Barandiaran, Egbert 2014: 5).

To explain this point more in depth, it is worth to have a look at the primary sources that enactivists consider in their explanation of the cognitive relation between organism and environment. As previously pointed out, the enactive approach to cognition describes the environment in terms of Umwelt. This is a concept developed by the Estonian theoretical biologist Jakob Von Uexküll (*Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten* - 1934), and that has been widely used in Merleau-Ponty's thought (see in particular the series of lectures given at the Collège the France in the late 1950's)[20], in which enactivists find the source of many of their ideas.

---

[19] Within the philosophical frame in which this discussion develops philosophers usually talks about mental states only, because they are considered to be the locus of cognition. Nevertheless, in my discussion, I extend the argumentation to behaviors and actions. Indeed, as implicitly shown in this paragraph, the enactivist perspective does not endorse the classical framework of philosophy of mind (folk psychology, functionalism): rather than identifying cognition with the ownership of mental states (that can be accessible or non-accessible to consciousness) it describes cognition as the *possibility of acting meaningfully* (see the concept of enaction).

[20] See Merleau-Ponty, M. (2003) *Nature: Course Notes from the College de France*, D. Seglard [Ed.], English translation by R Vallier, Evanston, Northwestern University Press.

The central idea of Von Uexküll's biology consists in claiming that each living being lives in its own Umwelt (built by means of its actions, behaviors, perceptions and so on), which is considered to be a subjective universe. It can be described as a habitat, namely a world that is not made by sets of neutral objects, but that is rather a domain of "performance qualities" (Uexküll 1982: 28), namely objects that are "subjectively tuned" because they *afford* some action and prevent the agent from doing other actions.

To explain this I consider an example developed by Uexküll (Uexküll 1982: 27). An angry dog barks at me while I am walking in the street. In order to drive it off, I frighten it by grabbing a stone and throwing it in the middle of the street. To the dog, the stone I throw at him had acquired a 'throw-quality', and this quality makes sense to it. The boundaries of its Umwelt include the perception of this performance quality. That is to say, the animal is able to respond meaningfully to this quality: according to the context, it can run to grab the object that has been thrown (if the context is playful), or it can run away in order to rescue itself (if the situation is perceived as a dangerous one).

Nobody observed my action, and the stone is left in the middle of the street. The time runs and the stone is incorporated in the country road: it will serve as a support for the walker's street. The shape of the stone, its weight, its physical and chemical properties have not been altered. Nevertheless, something really relevant changed: its meaning. To the walkers that enjoy a walk in this countryside street, the stone had acquired a 'path-quality'. It has now a meaning that it did not have in the dog's perspective. Indeed, having a walk following a defined path delimited by stones is something that humans do, animals usually do not.

Even if this example is very simple, it shows one of the most important ideas at the core of Uexküll theoretical biology, and it can be helpful to make clear an epistemological point of the enactivist perspective on cognition.

Within the boundaries of an Umwelt, objects play the role of *meaning-carriers*: they invite the living being to perform actions that can be judged as meaningful because they are occurrences of the comprehensive *meaning-plan* (Uexküll 1982: 43) of a living creature. The concept of Umwelt refers to specie-specific meaning-plans: each specie builds a meaningful universe according to its needs and possibilities, and it does that according the possibilities of action its bodily and motor structure offers.

The existence of meaning-plans (that is *motivated* domains of meanings), which should be interpreted as the horizon of possible actions that an agent can perform, is precisely what provides the *reasons* according to which actions performed in the world are accountable for cognitive judgments.

Sense-making provides an *internally consistent space of meanings* that allows the interactions between organism and environment to be judged: i) as meaningful (and according to the enactivist approach this meaningfulness is precisely what cognition is); ii) as fulfilling or not fulfilling what the overall meaning-plan requires. It is precisely in this sense that it has been claimed that sense-making endows the world with a normative status: it sets a ground of related meanings the organism brings forth or do not bring forth while it acts. In this sense it can be said that there is a criterion to judge if an action is successful or not for the preservation of life, if it consistent with the life-meaning-plan of the organism. The possibility of being "wrong" (in this case this means performing an action that does not match the meaning-plan of an organism, and then "wrong" means "unsuccessful") lays the foundations of cognition, or better, of cognitive behavior, by distinguishing it from a mere reaction. It situates actions in the realm of reasons and not in that of causes.[21]

This point about the way the enactivist approach to cognition looks at the normative domain the possibility of cognition rests on is particularly interesting for the purposes of this research, because it not only determines a radical distinction from classical views of cognitive sciences, but it also distinguishes enactivism from other theories of the "4 E Cognition Debate".

I briefly explained why normativity is a necessary condition to talk about cognition. If there is a normative domain one can refer to, there is also the possibility to make what -according to the rules that define the normative domain- is a mistake. This constitutive connection between cognition and the possibility of being wrong goes back to Descartes, and it shaped the overall discussion in philosophy of mind until nowadays. Usually the issue of the relation between cognition and normativity is faced by appealing to the concept of *representation.* Indeed, the apparatus of mental representations has been developed in order to account for the *epistemic gap* between mental contents (which show themselves according to reasons) and contents of the world, namely "factual information". As previously explained, the contents of the mind, in contrast to those belonging from world, might or might not be what they should be (Rowlands 2012: 134). To account for this possibility of error, mainstream approaches to philosophy of mind conceive of mental states as *relations to representations.*

[21] In this part of my argumentation I implicitly make use of some ideas of the debate on "spaces of reasons" and "spaces of causes" developed in Sellars' book *Empiricism and the Philosophy of Mind* (in particular §36), and then in McDowell's *Mind and World.*

Even if there is not a unitary perspective on representations within the contemporary debate in Cognitive Sciences, it seems to be the case that different theories usually share a core point about the epistemic efficacy of representations. Representations can actually account for the normative dimension that lays the foundations of the realm of cognition because they are discrete items (tokens) that can be tokened incorrectly (Rowlands 2006a: 167): representations can *misrepresent* what they are supposed to stand for. The relation of mediation they entertain with stimuli of the world (what is supposed to cause it) can be correct or incorrect.[22] Therefore, by conceiving of mental states as relations to representations, those approaches to philosophy of mind and cognitive sciences accounted for the possibility of mental states to be wrong: their content can misrepresent external reality, and then it can be judged as wrong/incorrect/false.

Now, what is interesting about enactivism is the fact that this philosophical stance takes into account the issue of normativity, cognition and error in a way that is radically different from those standard approaches to philosophy of mind. In my discussion of this problem according to the enactive approach, the notion of representation has never showed up. Indeed, the picture of cognition the enactivist perspective draws is *non-representational*. The concept of representation is completely substituted by that of action (and this means many things: perceptually guided action, enaction, sense-making, skillful action, intelligent behavior, and so forth).

This shift from representation to action is strictly connected to the idea of life-mind continuity, to an experiential conception of cognition, to a strong embodied perspective, and to an explanation of meaning that sees its origins in Uexküll theoretical biology. This can be expressed with a citation from Varela, Thompson and Rosch, which sums up great part of the discussion offered in this paragraph.

> "The key point is that such systems [= living organisms] do not operate by representation. Instead of *representing* an independent world, they enact a world as a domain of distinctions that is inseparable from the structure embodied by the cognitive system" (Varela, Thompson, Rosch 1991: 140, text in brackets added)

[22] In this paragraph I do not take into account the relation between the normativity condition and the misrepresentation one in detail, and I do not even talk about other "representational conditions". A detailed exam of this problem will be developed in Chapter II. For the moment, I aim to discuss just some general points, which are useful to understand some of the core ideas of the enactive approach to cognition.

The citation says that living organisms are tied to their Umwelts because the "historicity" of enaction (that is the effects of repeated transactions between organism and environment, which endow the subject with a practical knowledge with a "knowing how to act" or "being ready to do [something]" according to environmental circumstances – Noë 2004: 2, text into brackets added) depicts the agents' world as a "field of relevant affordances" (Rietveld, Kiverstein 2014; Rietveld 2008a; Rietveld 2012a) inseparable from the embodied structure of the agent. The agent and the Umwelt co-emerge as *structurally coupled* (Varela, Thompson, Rosch 1991: 151). This means that the cognitive agent and her world become a *unified system*, whose parts are interdependent and inter-defined.

To explain this point, I try to offer a thought experiment. Think of the activity of climbing, in which a climber has to cope with what in cognitive science would be described as a problem-solving task: heading towards the top of a rock by following the simplest and safer route. According to an enactive approach to the problem, the climber and the mountain would be described as a coupled system. On the one hand, the repeated actions of the climber, or better of the community of climbers, modified the structure of the rock, and this made the physical structure of the mountain suitable for the performance of successful practices. Not only the rock is full of spits for the passage of ropes, but also it is quasi-immediate to guess the best route to follow, because the climbers' repeated actions left traces of their activities on the rock. For example, the color of some parts of the rock is lighter because its material surface has been repeatedly rubbed by the climbers' feet and hands. On the other hand, the trained climber's body has changed in virtue of the repeated interaction with the rocks. For instance, professional climbers' hands become a bit curve, and the texture of their skin harder because they use to spend a lot time holding parts of the rock, keeping their hands in the grabbing/holding position. The interaction between the climber and the rock created a system that makes cognitive practices successful. The modification of the climbers' bodily structure (and their knowing how to act) and the modification of the rock, makes the process of problem solving -which is identified in the on-going embodied action of following the best route and not in a speculative conscious decision[23]- economical in terms of time and physical (and emotional) effort.

[23] This problem-solving task has been described in this way, namely by stressing on implicit practical knowledge, in order to be faithful to the enactivist approach to cognition, which de-emphasizes intellectualist aspects of cognition. It is also worth to notice that, concerning the example offered here, I agree with this non-intellectualist account of problem-solving. Empirical evidence that supports this claim can be found, for example, in Milner and

The success of the cognitive activity depends on the overall dynamics of world-engaging practices, in which human actions and environmental modifications are co-attuned.

Looking at this example, it can be explained how, within the enactive explanatory frame, one can account for "errors", which, as already explained, are fundamental for the definition of cognition. Those ones would not be explained in terms of misrepresentation of the external reality, but as gaps in the continuous dynamical coupling between the cognitive agent and her environment, as breaking points in the habit-based transactions between parts of the system.

As already explained, the climber is endowed with a multilayered implicit or unreflective knowledge about her practices. As any embodied agent, she knows how the perceptual array will change according to her movements (sensorimotor dependencies), then she will move her body in a certain way in order to accommodate her "perceptual expectations". Moreover, since she has been trained, she is endowed with an immediate responsiveness to the rock shapes: those ones are not just shapes, but they are holds, surfaces perceived according to the "grabbing" or "holding" modality. Furthermore, she is sensitive to the texture and to the color of the rocks. According to the climber-rock coupled system, this means being sensitive to the viability of a route.

Focusing on the "skilled" level of the climber's embodied practical knowledge, it can be said that errors can be explained in terms of a temporary disattunement (Bruineberg, Rietveld 2014: 9) between the agent and her environment. For instance, suppose that our climber goes to train herself in a new place. The rock wall looks exactly as the other rock walls she saw in the past, excepting for a detail she does not know. For some reason, non-viable parts of the rock have been artificially rubbed and de-colored and they look like what, according to the climber's implicit practical knowledge, would be a track. Obviously, she does not have an accurate knowledge of the rock-wall structure. It is the first time she goes there, so she cannot recall some memories about the structure of the wall. Moreover, even if she looks at the wall, she cannot have an accurate image of what she will find while she will be engaged in the activity of climbing, because the perceptual experience of the rock wall

---

Goodale's neuropsychological research on visual perception (which some works that endorse a radical enactivist approach to cognition take into account; Hutto, Myin 2013: 46-50). This research shows that agents can perform complex manual acts in a very accurate way even when they lack the capacity to give verbal reports of their activities or to explicitly describe the visual scenes they are engaged with (Milner and Goodale 1995: 128 -138, on subjects affected by visual agnosia).

is a paradigmatic and exaggerated example of a fundamental feature of perceptual arrays: they are deeply perspectival and they vary according to the perceiver's bodily postures and movements. Looking at the wall from the bottom of it, she cannot figure out what she will have to cope with.

The climber starts to climb and she heads to the point where those "fake holds" has been placed. She reacts to the present environmental conditions as she did in the past in similar circumstances. Her practical and habit-based knowledge about climbing makes her ready to move her body in a certain way (she co-ordinates her movements, she moves parts of her body at a certain speed, and so on), acting upon the context by following some implicit embodied practical rules. Nevertheless, while performing what her "knowing how" suggests, she fails to cope with those sabotaged pieces of rocks. Probably, a supporter of representations would explain the climber's failure by saying that a misrepresentation of stimuli from the world occurred in her mind. She did not represent the structure of the rock accurately: since her mental state of believing was related to a wrong representation, her belief that that piece of rock was a reliable hold was false.

An enactivist would rather explain this case of error by saying that she acted relying on what her body learnt during her recurrent previous experiences with rocks, she immediately deployed a contextual set of skills which worked well in past engagements with the environment-rock, but she acted in this way within the boundaries of the wrong system.

She relied on embodied skills that experience shown as reliable in the past, but that were not attuned with that (partially) new practical context. The error that occurred within that situation then depended on a temporary *decouplement* or *disattunement* between the agent and her environment. Something was perceived as an affordance for a specific action, which become quasi-stereotyped thanks to recurrent habitual practices, when the context required a different kind of practical arrangement.

The cognitive error occurred because of a lack of context-switch[24]: this temporarily turned the skillful agent into a stranger in her own land of affordances. This is to say that cognitive errors, rather than happening in the head, namely because of the occurrence of a non-accurate or wrong representation a mental state is related to, occur when the "inter" of inter-actions between the agent and the environment is filled with a wrong relation,

[24] For a detailed account of the relation between cognitive efficacy and context-switching in a non-representational perspective see Rietveld 2012a.

which produces a temporary crack in the dynamical coupling that habitually makes the organism and the environment a persistent unity of processes.

I.8 Embedded Cognition: discussing the project of locating cognition in externalities.

In the previous paragraph, the project of extending the boundaries of the mind made a step beyond. The enactivist approach to cognition, focusing on the idea of biological embodiment and on phenomenology, describes cognition as a dynamical process that takes place in the interaction between the embodied agent and the Umwelt. This entails a relational account of the mind. The realm of the mental is identified with the continuous inter-relation of processes, whose locus cannot be actually individuated neither in the cognitive agent's singularity (i.e. in her brain, in her body), nor in discrete informational occurrences the context of a cognitive practice offers.[25] The mind is rather explained as the relation between the agent and the environment, interaction that is realized within the dance of embodied actions, their effects on the environment, and the effects that these meaningful and cognitively relevant parts on the world have on the deployment of action. As shown in my discussion, this approach to cognition is more radical than other embodied explanations of cognitive experiences, because it really stresses on the way our bodies are tied to the context of cognitive practices. More precisely, it can be said that the contexts of cognitive practices cease to be just locating factors of cognition, becoming what can be defined a *cognitive agent*. In this way, the world, in its being cognitive, becomes a part of the mind.

Another way to include what is external to the boundaries of the skull and of the cognitive agent's biological body into the realm of the mental is offered by what is known as "Embedded Cognition". To be accurate, it is worth to notice that, within the contemporary debate about "4 Es" approaches to cognitive sciences, the label "Embedded Cognition" is not used in a univocal way. Sometimes it is a synonym of "situated cognition", expression that refers to the broad project of re-situating cognition in the real-world environment (Wilson 2002; Dawson 2014; Rupert 2009), and this claim is endorsed by all the philosophical approaches of "4Es", at different degrees and by developing different philosophical strategies. Moreover, the expression "embedded" is sometimes coupled with "embodied" (e.g. Clark 1999; Stapleton, Ward 2012), and it is used to describe how cognition unfolds through the coupling between the organism and the environment, as I tried to explain in the previous paragraph considering the enactive approach to cognitive sciences.

[25] For this reason the enactivist stance has sometimes been said to be the "*tertium non exclusus*" in the debate between internalists and externalists (Thompson, Stapleton 2009).

Generally speaking, the phrase "embedded cognition" has been coined to stress the fact that cognition is a practice that unfolds in a given natural and cultural context. This last one functions i) as a constraint for cognitive practices (namely it affords some actions and it prevents the agent form doing other actions – see approaches that widely rely on affordances, e.g. Rietveld 2012a; Rietveld 2008a; Rietveld, Kiverstein 2014), ii) as an enabling condition and as a constitutive factor (namely it makes the realization of some kinds of contents possible– Clark 2008; Clark, Chalmers 1998), iii) as a form external scaffolding (namely as a reliable and trustable support for the management of the cognitive load– see Sterelny 2010; Clark 2003; Clark 2005a; Clark 1998).

Nevertheless, despite it is true that the broad project of "4Es" points to the situated nature of cognitive process, it seems to me that this broad use of the term "embedded" is misleading. "To embed" means "to fix firmly in a surrounding mass of some solid material", and the word was coined in the domain of geology to refer to fossils in rocks (Oxford English Dictionary: http://www.oed.com/view/Entry/60835?redirectedFrom=embed#eid). Hence, if the etymology of the word "embedded" is taken into account seriously, the term should be used to refer to something that is *firmly located* in something else. Therefore, relying on this original meaning of the word, I suggest that the phrase "Embedded Cognition" should be used to refer only to those theories which endorse a "locational claim" about cognition (i.e. cognition is located in the environment or in parts of it). That is why I prefer not labeling theories such as the enactive one as "Embedded Cognition", because Enactivism rejects any "locational claim" about cognition and it widely relies on the concepts of process and dynamicity, features the verb "to embed" cannot account for.

I rather prefer using the expression "Embedded Cognition" to refer to the claim that cognitive processes entertain a relation of *dependence* with environmental structures (Rowlands 2010: 69). This is to say, Embedded Cognition claims that, while accomplishing some cognitive tasks, a cognitive agent manipulates parts of the environment in a successful way, namely in such a way that the cognitive work she does internally[26] is made lighter. In this way, the cognitive load (or "epistemic credit", to quote Clark and Chalmers – Clark, Chalmers 1998: 8) the task requires is distributed between internal resources and the environmental ones. Then, the basic idea of Embedded Cognition consists in stating that some cognitive processes should be considered as heavily dependent on environmental structures because in absence of the appropriate

[26] I.e. mapping the environment by means of mental representations, mentally planning the steps of the cognitive activity, building internal predictive images of the results of the cognitive practice, and so forth.

environmental structures the agent may not be able to successfully accomplish that task (Rowlands 2010: 69), where when I talk about success I refer to the speed and the accuracy of the task (Kirsh 2010: 443). Then, since the realization and the success of some cognitive task is strictly dependent on how things are in the environment, Embedded cognition claims that, in order to understand how cognitive processes deploy, the scientist should individuate those discrete parts of the environment (that is external cognitive structures) that play this relevant function in cognitive processes. The context in which a cognitive task is performed is included in the picture of cognition the Embedded perspective draws because parts of it function as the *locus* in which some cognitive work is get done.

An example useful to explain what Embedded Cognition is can be found in Kirsh and Maglio's work, in particular in the article "On distinguishing Epistemic from Pragmatic Action" (Kirsh, Maglio 1994). For the sake of accuracy, it is worth to notice that these two cognitive scientists do not make use of the phrase "Embedded Cognition". Nonetheless, I claim that some of the theoretical insights this article offers, and also the experimental results the two cognitive scientists consider, are consistent with the picture of "Embedded Cognition" I want to draw. Moreover, the case study the two scientists present is particularly interesting for the consideration of a philosophical point that, as it will be shown along the development of my research, makes the difference in the contemporary debate about "4Es": computation.

The article I am taking into account aims at drawing a distinction between pragmatic actions and epistemic actions. The first kind of action is identified with those sets of movements whose function is to bring the agent close to her physical goal. An example of pragmatic action is what a football player does in order to reach the ball that runs on the grass: she moves her body in certain way, she implicitly checks the balance between different bodily parts, etc. The task the agent is supposed to perform has a mere physical nature: the agent interacts with her environment in order to bring forth physical conditions that satisfy her goals. The second kind of action, the epistemic one, is a physical action as well, but its effects do not just lay in the realm of facts, to use the distinction considered in the previous paragraph. Differently, the term "epistemic action" refers to those actions that, by producing physical effects in the environment, modify the agent's computational states. Epistemic actions are external actions the quality of the cognitive the agent's internal states (i.e. computational states) *depends on*. They are actions performed upon some parts of the environment; those parts of the environment, when manipulated

in a certain way, function as *external representations,* or as external cognitive structures, on which a computational work is done.

That is why the quality of the computational states internal to the cognitive agent changes: they become "lighter" because some computational work is get done outside, in the world. To explain this point in another way, it can be said that epistemic actions are physical actions that, in their making *information* easily available by producing changes in external structures, simplify the cognitive agent's tasks because they

i) reduce memory-work involved in mental computation;
ii) reduce the number of steps involved in mental computation;
iii) reduce the probability of error of mental computation (Kirsh, Maglio 1994: 513).

To verify this hypothesis about the way external structures produce a distribution of the cognitive load of a task, Kirsh and Maglio focus on the case study of Tetris, considered as a paradigmatic example of what cognitive agents do when they cope with real-time problem solving tasks performed in the real world.

Tetris is a real-time interactive video game in which the player has to maneuver falling zoids of different shapes into specific arrangements on a screen. The zoids fall from the top of the screen one at time, and each zoid keeps on falling until it reaches the bottom of the screen or the surface of other zoids that previously landed. When a row of zoids is built, this arrangement of shapes disappears from the screen. The aim of the game is to keep the screen as clean as possible: when Tetris bricks reach the bottom of the screen the game is over. In order to keep on playing, the player has to build rows of zoids, maneuvering the zoids as quick and accurately as possible.

To account for successful problem-solving tasks of this kind, the two cognitive scientists individuated two possible explanatory models.

The classical one is called "process model". It relies on classical cognitivism's assumptions, coupled with those of good old-fashioned artificial intelligence (GOFAI). This approach would explain the problem-solving task of Tetris playing by saying that the cognitive agent internally manipulates representations of the Tetris zoids, computing the best place to put the zoids and the best trajectory of moves to place them. This is to say, according to a classical information-processing of Tetris Cognition, each action performed in the real-time game playing (e.g. motor control) is forerun by an internal planning of action, described in terms of computations performed on internal

symbolic representations. Indeed iconic representations of the Tetris bricks are said to be encoded in symbolic representations, according to a process similar to that of transduction described in §I.5. Hence, according to this model, the player is said being able to successfully perform the Tetris problem-solving task because she builds an internal map of the external representational structure the game is made of. She is said to have a cognitive grip on the structure of the cognitive task, and consequently to perform a successful set of actions, because she re-works the external cognitive situation, by internally planning motor action.

The second model is built on the "epistemic action claim". The basic idea of this model is that the player is able to successfully perform the cognitive task because she is trained to fluently maneuver (rotate or translate, where by "translating" I mean moving left to right or in the opposite direction) the zoids on the screen using a joystick. The player is said to perform the problem-solving task without following a plan embedded in her working memory; she rather relies on the external representations available in the action-context of the task, making the information they bear easily available by performing epistemic actions. While the agent moves the zoids on the screen, she modifies the physical structure of a portion of the environment, and in doing so she modifies the informational states the cognitive task implies. By means of action, she actively alters her perceptual domain, and within this alteration, information relevant for the accomplishment of the problem solving-task (i.e. spatial relations among the zoids, compatibility between shapes) is immediately made present in the environment. The success of the cognitive practice is said to depend on this "cooperative and interactional relation with the world" (Kirsh, Maglio 1994: 546): problem-solving and decision-making are constitutively dependent on *(epistemic) action-perception loops*. They entertain a very tight relation of dependence with what is outside of the subject's head: some cognitive operations are directly performed on the screen where representations are embedded.

In order to test this two explanatory models, namely i) the idea that problem-solving cognition takes place in the head versus ii) the idea that those kinds of cognitive tasks are performed by distributing the cognitive load between internal and external cognitive structures, Kirsh and Maglio made an experimental study in which two cognitive strategies, expressive of the two explanatory models, were confronted. i) They collected tachistoscopic tests of subjects that performed mental rotation tasks related to Tetris, they implemented a program called "RoboTetris", built on classical information-processing model of game expertise, ii) and they observed the behavior of

agents engaged in real-time Tetris playing, focusing on two kinds of actions performed while manipulating the joystick (rotation and translation), recording the timing of keystrokes.

What emerged from the comparison of these experimental situations is that neither the results of tests on RoboTetris, nor those emerged from mental rotation tasks were consistent with those from the observation of an embodied agent playing Tetris. For instance, according to a classical informational-processing model of Tetris cognition, the bigger the time window in which a subject can plan her moves (relying on mental rotation of zoids) is, the fewer the external manipulations of Tetris bricks are. Experimental results contradicted this hypothesis, showing that rotations and translations acted on the screen occur in abundance (graphs of those experimental results, and a wider discussion of data, can be found in Kirsh, Maglio 1994: 523 – 524).

According to Kirsh and Maglio, those results confirm their hypothesis: agents that are well adapted with the environment where a cognitive task takes place know how to balance internal and external computation. That is, subjects are able to successfully perform a problem-solving task relying on a variety of cognitive strategies which, in some cases, rely more on what is present in the context of the task and on what can be done in the environment rather than on the intellectualization or internalization of what is in the environment.

In this sense, some cognitive practices have an *embedded nature*: their success and quality depend on external actions (epistemic actions) that make the external world being *its own representation*, to explain this point by making use of Brooks' words (Brooks 1991:140). Epistemic actions, by successfully manipulating information available in the environment, make the cognitive agent able to acquire a grip on the situation she faces because the relations between parts of the environment (i.e. zoids of different shapes, concave and convex spaces, and so on), on which the success of the cognitive practice heavily depends, are externally represented.

Now, after this very brief discussion of Kirsh and Maglio's work, I would like to focus on some core points of this approach to cognitive practices, which, on the one hand distinguish this explanation from other philosophical stances (i.e. Embodied Enacted Cognition), on the other hand lay the theoretical foundations of the Extended Mind Hypothesis (EMH).

In my explanation of Embedded Cognition, I focused on the concept of epistemic action. This attention to action in cognition, the realization of a cognitive practice is said to be strictly connected to, can make the reader thinking that this approach to cognition is totally consistent with the enactive

one. Indeed, as explained in §I.7, Enactivism is grounded on the philosophical claim that cognition *is* action (Varela, Thompson, Rosch 1991), and that cognition (namely knowledge considered in its active and operational aspect[27]) constitutively depends on an implicit, embodied practical knowledge (Noë 2004). To me, this similarity has a superficial nature only. Indeed, even if Kirsh and Maglio claim that their approach accounts for a "cooperative and interactional relation with the world" (Kirsh, Maglio 1994: 546), aspect that would be consistent with the relational account of cognition Enactivism points to, it seems to me that Embedded Cognition underplays the wider implications of *inter*-action in cognition. If it is true that on the one hand this distribution of the cognitive load between internal and external representations seems to point to the "*inter*" of interaction, on the other hand, focusing on the concept of *computation,* this Embedded Account of cognition seems to miss the *relational* aspect of cognition it wants to account for.

Indeed, in order to describe cognition as a relational process, the scientist should look at what is in the middle of the *relata*. To explain this point in another way, it can be said that to account for cognition as a relation, a theory should conceive of cognition as the dynamical process by means of which two elements become related. Enactivism seems to satisfy this philosophical requirement because it focuses on the concept of sense-making, which makes the organism and its Umwelt emerging as interdependent through action. The Embedded approach to cognition (as taken into account in this paragraph) seems to miss this point[28]. Even if Kirsh and Maglio consider problem-solving as dependent on action-perception loops (where the concept of loop would account for process and dynamicity by pointing to the circularity of action and perception), the two scientists tend to depict action as *isolated* and *discrete*. They do not take into account action in its dynamical and processual development; they prefer focusing on the *passages*, on the *operational steps* in which external representations are manipulated. This is dependent on the implicit endorsement of *wide computationalism* as an explanatory frame. By "wide computationalism" I refer to an approach to cognition that reassesses the core ideas of folk-psychology (which individuates cognitive states in computational states) by claiming that computational states, namely cognitive states, do "not supervene on the intrinsic, physical states of the individual"

[27] For this definition of cognition, which accounts for the etymology of this word (form the Latin *cognitionem*: getting to know, ability to know), see Steiner 2008: 86.

[28] My objection concerns the article published in 1994 only. Indeed, in other publications (e.g. Kirsh 2010) a more relational account of cognition is given by re-working on the enactive concept of sense-making.

(Wilson 1994: 352), but they can be instantiated also by what is outside of the individual subject. Broadly speaking, computations are "effective procedures" (Aizawa 2010: 227), procedures whose states and rules can be made explicit, and that are performed on representations. In the case study of Tetris, zoids can be considered as external representations (they stand for the bricks of an imaginary wall) that are computed, that is effectively manipulated, externally, making the information embedded in external representations available.

The way Embedded Cognition accounts for action is peculiar: epistemic actions are steps of a computational processes, they are operations that can be individuated as *discrete* because they show, make available, informational states (e.g. the translation from right to left, performed in one motion, can light up the relation between a zoid and the space where it is supposed to land). Action is modeled on the concept of computation. In this sense, even if Embedded Cognition accounts for the dependence of cognitive processes on what is beyond the boundaries of the individual, the philosophical framework it relies on is still classical, conservative, I would say. It endorses an explanation of the relation between information, representation and computation that reproduces in the external world what, according to classical cognitive psychology, happens within the boundaries of human mind, by saying that the contents of external representations are made "cognitively manageable" via computations.

Moreover, this approach to cognitive practices, even if it accounts for the dependence of cognition on external structures, still relies on a classical account of what happens in the human mind when a cognitive agent is engaged in a cognitive practice. Indeed, cognition is said to be embedded in the environment to the extent that computations on discrete parts of the environment alter the subject's internal computational states. The first term of the comparison between internal and external cognitive processes is "computation on mental representations". Then, the appraisal of a cognitive process as partially embedded is dependent on the appraisal of changes in internal computational states.

For those reasons, the Embedded approach to cognition is appreciated by philosophical stances which endorse the claim that cognition is dependent on external structures, but that reject the idea that the mind is extended in the environment (Rowlands 2010: 70, which refers to Rupert 2004; Adams, Aizawa 2001; Adams, Aizawa 2009; Adams, Aizawa 2010b). This makes the position that Embedded Cognition occupies in the debate of "4Es" even more unclear than what was shown in my terminological discussion of the phrase "Embedded Cognition". On the one hand the fact that Embedded Cognition i) endorses a

dependence claim, ii) and develops its argumentation relying on the assumption that the criteria to define the "embeddedness" cognition should be found in the appraisal of the cognitive agent's internal states modifications *caused* by the engagement with external structure, make this approach consistent with theories that individuate the mark of the mental within the boundaries of the individual. On the other hand some concepts of this approach (e.g. external representations, wide computation, external manipulation of information, epistemic action) are reassessed by the Extended Mind Hypothesis, which not only claims that cognition is *dependent* on or *caused* by external structures, but also aims to account for the cases in which the Mind itself extends in the environment (*constitution* claim).

I.9 The Extended Mind. A radical theory to extend the mind?

As sketched out in the previous paragraph, the Extended Mind Hypothesis (EMH) claims that, in some cases, cognition extends (namely it is *realized* by) into parts of the world. Then, if the word "mind" is used to refer to cognitive processes, in those said cases of cognitive extension, the mind should be said to extend (namely being *constituted* of) into the world.

To explain this thesis, EMH, in Clark and Chalmers' version and in its following developments by Clark (Clark, Chalmers 1998; Clark 2008), begins its argumentation by endorsing a philosophical position about the extension of cognition and the mind that has been defined as content-enabling externalism (Hurley 2010: 105). According to Hurley's taxonomy of externalisms, *content-enabling externalism,* or *vehicle-externalism* or *active externalism* (Clark, Chalmers 1998 for this last expression) is a peculiar kind of *how-externalism*. How-externalism is a philosophical stance that is more radical than classical externalist theories about mental contents. Classical externalist theories of mental contents claim that the content of mental states is determined by the external natural world (e.g. Putnam 1973) and social world (e.g. Burge 1979). How-externalism claims that not only intentional contents of the mind are dependent on the external world, but it also claims that the processes, mechanisms or vehicles mental states are enabled by extend beyond the boundaries of the human head. This is to say that not only the "what" of mental states is determined by the external world (e.g. the content of mental states about water is determined by the history of interactions[29] with water) but also the "how" of mental contents, that is the way mental states unfold, is sometimes external. The *material carriers* of mental contents cross the boundaries of human heads, spreading out "across brain, body and certain aspects of the physical environment itself" (Clark 2005b: 1).

This claim about cognitive extension can be unpacked by considering the very famous (and controversial) thought experiment of Otto and Inga, presented in the seminal paper "The Extended Mind" (Clark, Chalmers 1998: 12-16).

This thought experiment is about the cognitive practice of two subjects. Inga and Otto want to go to an exhibition at the Museum of Modern Art in New York (MoMA) and one day decide to go for it. To fulfill her desire, Inga thinks

[29] According to EMH, this focus on the historicity of interactions with the external world makes this position about the acquisition of mental contents a form of passive externalism: parts of the world present in the real-time engagement of the agent with her environment play no significant role in driving on-line cognitive processes (Clark, Chalmers 1998: 8).

about the location of the museum, she remembers that it is in the 53rd street, and she goes there. To reach her goal, she just relied on her belief that the MoMa is in the 53rd street, caused by the information about the location of the museum embedded in her long term memory.

Otto, who suffers from Alzheimer's disease, has to adopt another cognitive strategy. Indeed, he cannot recall information about the location of the museum embedded in his long-term memory because this cognitive function is partially damaged. To overcome the problems his disease entails, he progressively developed a strategy of cognitive balance (or substitution). He always carries a notebook with him. In this notebook, he writes down new information when he learns it. When he needs to have access to information learnt in the past, he checks his notebook. Then, when Otto has to recall the location of the MoMa, he looks at the address he previously wrote down in his notebook, and he goes to the museum.

The two cognitive agents successfully performed the same cognitive task, even if they followed different cognitive strategies. In Otto's case, the material vehicle of information was not embedded in his head, but in an external resource. According to the EMH, this is a case of "cultural extension" (Hurley 2010: 127) of cognition. By relying on external resources explicitly designed for the agent's engagement in a cognitive task, this last one can be judged as successful (i.e. the cognitive agent reaches her goal).

Explained in this way, Otto's example seems to point to the same theoretical direction taken into account in previous paragraph, where I discussed the Embedded approach to cognition: the success of a cognitive task heavily *depends on* external structures or external representations. Nevertheless, the EMH wants to go a step further. Indeed, this thought experiment was formulated in order to demonstrate a more radical claim: that about the *constitution* of cognitive systems. Contrary to what described in the previous paragraph, in which an ontic perspective was endorsed (namely the aim of the theory was explaining "how the things are"; Rowlands 2010: 56), the EMH endorses an *epistemic claim*, which has *ontological consequences* about cognition and the mind.

What this thought experiment teaches is that cognitive practices does not just depend on external structures, but they are also built on them, they are constituted of them. This point is explained by Clark and Chalmers by means of the concept of *coupled system*. In the 1998 article, the concept of coupled system is explained in the following way:

> "In this cases, [namely those of "cultural extension"] the human organism is linked with an external entity in a two-way interaction, creating a coupled system that can be seen as a cognitive system in its own right. All the components in the system play an active causal role, and they jointly govern behavior in the same sort of way that cognition usually does. If we remove the external component the system's behavioral competence will drop, just as it would if we removed parts of its brain. Our thesis is that this sort of coupled process counts equally well as cognitive processes, whether or not it is wholly in the head" (Clark, Chalmers 1998: 9, text into brackets added).

A coupled system is a system in which a human cognitive agent and external reality, or better, parts of it, are linked by a two-way interaction. This means that the cognitive agent actively manipulates external parts of the world (i.e. she performs epistemic actions in order to retrieve relevant information), and the world, modified by epistemic actions, causes changes in the cognitive agent, namely it produces cognitive states (i.e. mental states). The two-way interaction the EMH talks about is said not being "episodic". On the contrary, this reciprocal influence is *continuous*. This continuity of the interaction between parts of the system makes the changes within the system flowing from one part to the other (van Gelder, 1995: 373), linking the parts of the system in a *loop* (Clark 2003: 75), namely in a lasting circle of causes and effects relevant for the deployment of the cognitive system's activity (Clark 2008: 87).

All the components of the system, its neural and extra-neural parts, play an *active role*: they drive real-time cognitive process. This active, non-distal role played by externalities makes them parts of the cognitive system: if Otto loses his notebook, the unity of the cognitive system is broken, then the cognitive practice cannot take place, because the cognitive agent (that is the coupled system) is not actually present in the situation of the task.

Therefore when the EMH talks about coupled cognitive systems it is not just referring to the causal role externalities play in a cognitive practice, in which there is a "ready made" cognitive agent that manipulates parts of the world.

I think that objections against the EMH miss this point, namely that concerning the *formation of the structure* of a cognitive agent. Examples given to demonstrate that the EMH makes a fallacious inference from causal relations to constitutive relations (this objection is known as "coupling constitution fallacy"; Adams, Aizawa 2001; 2009; 2010a; 2010b; Aizawa 2010a; Rupert 2004), integrating in the cognitive system what (according to those philosophical stances) is not actually cognitive, presuppose situations like the following ones.

i) The expansion of a metallic strip in a thermostat, which is causally linked to a heater that regulates the temperature of the room in which the thermostat is located (Adams, Aizawa 2001: 56);
ii) the causal interaction between the kidneys and parts of the circulatory system in the process of alternation of the impurities of the blood (Adams, Aizawa 2001: 56);
iii) the role of the economic conditions in Germany to understand the Nazi Germany's invasion of Poland (Rupert 2004).

Critiques of the EMH say that, in the same way

i) the expansion of a bimetallic strip in a thermostat (which causes a modification in the thermostat, and then in the temperature of a room) is not a process that extends into the whole system,
ii) the fact that parts of the circulatory system have causal influences on the work of the kidneys does not allow us to claim that the alteration of the impurities of the blood occurs within the circulatory system,
iii) the fact that, to fully understand Nazi Germany's invasion of Poland, one should have information about the economic conditions in Germany does not make those economical conditions part of the historical event of the invasion,

the fact that Otto's notebook plays an active causal role in driving a cognitive practice is not a sufficient condition to consider it to be a constitutive part of an extended cognitive system.

All those examples are aimed at demonstrating that, in order to endorse EMH, considering informational and causal relations only is not enough. According to those critiques, it is not clear how the constitution claim can be derived from the active causal role played by externalities in cognitive process, in virtue of their informational content.

As explained previously, the problem with those objections to cognitive extension rests in the fact that they do not focus on the way an agent becomes cognitive in virtue of the integration of internal and external cognitive resources. The examples offered above are expressive of this way of reasoning. They presuppose that there is a *discrete* agent, or item (the bimetallic strip, the kidneys, the agent's understanding of Poland invasion) "cognitively implemented" by the encounter with an external carrier of information/contents (Menary 2006: 333). Then they claim that the causal

coupling with those external discrete resources or structures does not make the said discrete system extending beyond its already defined boundaries.

The problem with those critiques consists in a misunderstanding of the original EMH project. Indeed, even if the EMH commonly focuses on cases of real-time, real-world cognition, it seems to endorse a philosophical position according to which we are allowed to consider a cognitive system as extending beyond the boundaries of the individual when there is an *history of couplings,* which theoretically builds the "integrated cognitive agent". Even if Otto's example deals with a defined cognitive task, according to Clark, what matters for an explanation of extended cognition is not just the real-time causal interaction between parts of the system (i.e. Otto and his notebook), but it is this real-time causal interaction *plus* the previous experiences of coupling that made the system *operationally* successful through time. According to the EMH, Otto and his notebook constitute an integrated system, and then, considering that discrete task, the cognitive practice is a case of cognitive extension, because

i) the information contained in the notebook was endorsed as true by the human agent in the past (and then it is considered to be trustable)
ii) and because the history of the engagements with the cultural artifact shown that the notebook is a reliable and easily accessible source of information (Clark, Chalmers 1998: 17; Clark 2010a: 84).

To put it in other words, according to the Extended Mind Thesis [30]supporters, and for "integrationists" about cognition (Menary 2006; Menary 2007; Menary 2010b) we can talk about extended cognition when a system that at given time built itself by coupling internal and external resources displays a cognitive behavior, which is made of hybrid processes. Then, the core claim of the EMH is not that cognition is extended because sometimes a pre-existing cognitive agent is coupled with external reality in a cognitive way. The core claim consists in saying that cognition extends into the world because some re-iterated episodes of coordination between a human agent and artifacts or external structures created something that did not exist before these cognitive engagement: an hybrid cognitive system (Menary 2006: 333-334, for the original version of this argumentative part).

[30] This is different from claiming for cases of cognitive scaffolding, namely cases in which, in virtue of causal relations, the environment functions as a scaffolding for a cognitive agent's practice (Clark 2008: 30).

Until this moment, my discussion of the EMH was focused on cognition. To be accurate, the argumentations offered here were aimed at making a philosophical hypothesis clear, the Extended Cognition Hypothesis (ECH), which, on the one hand lays the foundations of the EMH, on the other hand does not guarantee the extension of the mind by itself. Indeed, even if at this point of the argumentation the EMH made a step beyond Embedded approaches to cognition (which is compatible with a claim of mere dependence), claiming that some informational processes take place outside of the head is still compatible with an internalist account of the mind (i.e. truly mental states, such as beliefs and desires, are determined by states of the brain only; Clark, Chalmers 1998: 12).

To explain the way the EMH tries to overcome this narrow approach to the mental, by endorsing the ontological thesis of the partial external location of the mind, it is worth to look again at Otto and Inga's example, focusing on the relation between information and mental states.

The way Clark and Chalmers, and Clark in other publications (e.g. Clark 2008; Clark 2010a) explain how the mind is partially constituted by external structures relies on some ideas very common in philosophy of mind, namely functionalism about mental states. Functionalism claims that mental states should be defined in virtue of their causal and functional relations and not in virtue of their spatial location, or in virtue of the material substrate they are embedded in. Mental states do not necessarily supervene on the brain's physical constitution: being realized by a neural structure is not a necessary condition for the definition of mental states.

According to Clark's approach to the Extended Mind, claiming that mental states are necessarily embedded in human heads is not only consequence of the internalist prejudice that affected great part of the history of philosophy of mind, but it also depends on the unjustified premise that neural substrates are mental states because they are endowed with an *intrinsic content*. According to this approach to the mind, by conceiving of mental states as neural populations, a theory is allowed to think of mental states as entirely non-conventional carriers of a specific content. This would make them able to signify in a more direct way than conventional symbols do (Adams, Aizawa 2001: 48).

My point is that, not only the notion of intrinsic, or underived, content is obscure (and then what is left unexplained is the reason why we should think that intrinsic content is the mark of the mental), but also it is not clear how the intrinsic feature of mental content is tied to neural substrates. For instance, it seems possible to ascribe non-conventional contents to states of some artificial

creatures, or to silicon parts embedded in the brain, that do not have the same material constitution of neural areas (Clark 2005b: 4).

Given the vagueness and the lack of clarity that wraps the idea of intrinsic content, Clark and Chalmers claim that we should think of looser borders for the mark of the mental: it should be broadly identified with the function that an item or process plays in the *overall dynamics* of a cognitive practice in virtue of the *informational relations* that it entertains with other states, which carry other pieces of information.

According to this broad interpretation of functionalism about mental states, Otto and Inga's thought experiment can be unpacked in this way. Inga goes to the museum because she *believes* that the MoMa is on the 53rd street. The mental state of belief (interpreted as a disposition to act in a certain way in order to reach a goal) *causes* the intelligent behavior of going to the museum. Moreover, Inga believes that the MoMa is on the 53rd street because some *information* about the location of the museum is embedded in her long-term memory. The information embedded in her long-term memory is represented by the intentional content of the belief. The causal relation between information, mental states, and intelligent behavior has

i) an epistemic nature: it constitutes the explanation of the cognitive practice dynamics;
ii) it points to an ontological claim: if cognitive behavior (namely the effect of a cognitive practice) is explained by appealing to the relation between a mental state and the information it contains, then the mind (interpreted as a cognitive device) should be individuated in mental states that cause cognition.

Otto's cognitive situation can be explained by following the same argumentative strategy. Information about the location of the museum is contained in Otto's notebook, information causes the belief that the MoMa is on the 53rd street, and this dispositional belief causes intelligent behavior. As Inga's mind is constituted by her mental states, also Otto's mind is constituted by the same "functional substance". Nevertheless, in Otto's example, this functional substance is not entirely neutrally realized. Given that information represented as the intentional content of a belief is embedded in a cultural artifact, belief (namely a part of the mind) is embedded in an external structure as well.

To explain this point more accurately, it can be said that the relation that information, belief and cognitive behavior entertain in Inga's case is the same

as that entertained by information contained in Otto's notebook, his long-term memory belief and his behavior. Information contained in Otto's notebook plays the same functional and causal role of that one represented as the content of Inga's non-occurrent belief about the location of the Moma. Then, even if Otto's non-occurrent belief that the MoMa is on the 53rd street supervenes on a cultural artifact and not on brain structures, then the notebook should be considered to be part of the mental. Information it embeds causes the same disposition to act that Inga's internal belief causes: it has the same coarse-grained causal, functional, and explanatory role. Then, by endorsing a functionalist perspective on the mind, we do not have compelling reasons to let Otto's notebook outside of the realm of the mental.

Considering Otto and Inga's example, it seems to be the case that the argumentation for mind extension Clark and Chalmers give implies an inference from the *similarity* of function and causal powers to the ontological commitment of the partial external realization of the mind. If we reassess the thought experiment discussed here, it seems to be the case that Clark and Chalmers claim something like this: Otto's notebook is like an exogram. It is a set of memory records stored outside the nervous system (Sutton 2010: 189; Donald 2010: 71), which functions in a *similar* (or more successful[31]) way to engrams, namely memories embedded in the brain (Donald 2010: 71, who takes the expression from Lashley 1950). It provides information that can be retrieved along an extended time-scale; if incorporated in the cognitive agent's cognitive practice in the right way, it is a reliable cognitive medium; it drives a pattern of cognitive actions established through time, and so on. The notebook functions as a transparent equipment in the Heideggerian sense (Clark 2005b: 2). As Inga's engrams, or biological memories, it is implicitly "taken for granted" in the cognitive practice: thanks to its transparency, it fluently drives the subject's cognitive practices.

The appeal to the concept of similarity, towards which there have been many critiques, is confirmed by the first formulation of the Parity Principle, given in the 1998 article (Clark, Chalmers 1998). The Parity Principle says that

[31] For instance, we can have a "theoretically unlimited perceptual access to exographic records with various kinds of interface", they exceed life span, they can be retrieved by following different cognitive strategies, they offer a material substrate for flexible reformatting of memory records (Donald 2010: 72). This example concerning engrams and exograms, which seems to be useful to reassess Otto and Inga's example, gives me the chance to notice that the EMH (in its first formulation) focuses on the similarity between internal and external processes, not on the sameness of qualities or ways of information storage (Clark 2008: 115).

> "If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process" (Clark, Chalmers 1998: 8)

As the quotation clearly shows, according to the Parity Principle, which is one of the pillars of the first version of the EMH, the criteria to define something as mental presuppose internal structures or processes as the first term of comparison of this relation. As Di Paolo points out, even if the Parity Principle is aimed at laying the foundations of locational impartiality of cognitive processes, and then of the mind, the pre-theoretical understanding of what the mind is "is biased towards the inner" (Di Paolo 2009: 11), and, in the economy of a location neutral theory, this makes the appeal to the idea of Parity Principle a nonsense (Walter 2010: 288).

This objection, at least on a broad logical level, seems to really make sense. Even if other parts of Clark and Chalmers' argumentation provide interesting philosophical means to argue in favor of the external location of mental states, the appeal to this internalist pre-theoretical intuition about the mind seems to jeopardize the overall success of the general picture of the Extended Mind.

Nevertheless it is worth to notice that, in some of Clark's publications, the idea of Parity seems to be coupled with another notion, which does not entail this implicit primacy of internal structures: *complementarity*.

For instance, in *Being there* (Clark 1997a), book that lays the theoretical foundations of the EMH as described in 1998 article, he talks about "external structures [that] function so as to complement our individual cognitive profiles and to diffuse human reason across wider and wider social and physical networks" (Clark 1997a: 179, text into brackets added). Moreover he says that "The complementarity between the biological brain and its artifactual props and supports is thus enforced by revolutionary forces uniting user and artifact in a virtuous circle of mutual modulation" (Clark 1997a: 213). Furthermore, this idea of complementarity is endorsed also in works published after 1998. For instance, in 2001 Clark writes: "one useful way to understand the cognitive role of many of our self-created cognitive technologies is thus as affording *complementary* operations to those that come most naturally to biological brains" (Clark 2001: 20). In 2003 he says that "one quite general way to see the contribution of tools such as pen and paper is thus in terms of a deep complementarity between what the biological brain is naturally good at, and what the tool provides" (Clark 2003: 74). And again, in 2010, he says that "given sufficient complementarity and integration, (...) we may sometimes confront

hybrid systems displaying novel cognitive profiles that supervene on more than the biological components alone" (Clark 2010a: 99). Also the broad picture of the mind drawn in Clark's last book, *Surfing Uncertainty. Prediction, Action, and the Embodied Mind* (Clark 2016a), seems to convey this idea of complementarity. Indeed, all the discussion about the neural underpinnings of predictive/perceptual processes is integrated into an embodied and extended view of the mind, according to which "organismically salient (high precision) prediction error may thus be the all-purpose adhesive that, via its expressions in action, binds elements form brain, body and world into temporary problem-solving wholes" (Clark 2016a: 262).

This idea of complementarity, plus extended functionalism and the criteria for coupling, is particularly relevant to understand the pluralist approach to cognitive processes endorsed by Clark. The mind is said to be extended because cognitive processes take place in loops in which the cognitive functions of resources internal to the human head (i.e. representations, usually conceived as action-oriented representations or as predictive models; Clark 1997a; Clark 2016a) entertain a continuous exchange (i.e. are complemented by) with what is external to the head (i.e. functions and processes of the body, external representations, external computations, artifacts, and so on). Then, the explanatory tools this theory relies on are multiple, and this diversity of the explanatory tools is aimed at giving a multilayered account of mind processes.

Broadly speaking, I am sympathetic with the idea of extending the mind beyond the boundaries of human heads. I strongly agree with the idea (criticized by the EMH) that considering the mind is just as an internal representational device is a reductionist perspective, dependent on an internalist prejudice that, as shown by taking into account some objections to the EMH (i.e. the coupling-consitution fallacy, the intrinsic content requirement), seems to be not justified. I also think that this internalist prejudice depends on the implicit and arbitrary assumption, dominant along the history of western philosophy, that cognitive processes are processes by means of which the external reality is *mirrored* in the internal space of the mind, locus of cognition.[32] Nevertheless, I have some doubts about the way extended mind theorists account for this process of "mind-extension".

First, although the idea of complementarity sounds appealing, the multiplication of the explanatory notions it entails is not cognitively economic. In order to hold together different vehicles of information, and then of mental

[32] For this broad critique to the relation between mind and representations see Rorty 1979.

contents, the EMH relies on notions whose common ground is sometimes unclear. This problem will emerge in more detail in the next chapter, in which I will take into account how the EMH deals with the problem of representations.

Second, some aspects of the theory seem to be at odds with the general project of the EMH, namely that of giving a dynamical account of cognition, explaining it a set of *active processes* that cross the individual's boundaries (Clark 2008: 40 for this claim).

At a general level, it can be noticed that the very meaning of the label "Extended Mind" hides this problem. "Extended" points to something *static*, to the result of a movement that goes from "inside" to "outside" (and this was shown at another level of explanation by considering the objections to the Parity Principle[33]). To account for the process of mind extension, it is better to look at mind extend*ing,* namely to the process through which the very activity of the mind unfolds. To some extent, Clark's work points towards this direction. Indeed, the thought experiments and the cases that he takes into account deal with real-time and word-engaging cognitive practices. Nonetheless, it seems to me that the processual aspect the EMH aims to account for is still missed. This point is clearly taken into account by the enactivist critiques of the EMH, aimed to underline that the EMH, by endorsing a functionalist perspective, is unable to actually account for the *extensive* (Hutto, Kirchhoff, Myin 2014) nature of the mind.

For instance, Gallagher (Gallagher 2011b; Gallagher 2013) points out that, relying on the concept of (mental, representational, informational)*state*, the EMH implicitly endorses a static and *passive* explanation of mind, which is not consistent with the idea that the mind is made of activities, practices, and processes, claim endorsed by Clark at different levels (Clark 2008 and Clark 2016a). The functionalist frame in which the EMH is developed entails that, to find the "external mind", we should find the external *locus* of cognitive states (Di Paolo 2009; Paolucci 2011 for this objection). The problem with the endorsement of this locational claim[34] rests in the fact that it is at odds with a relational and interactive account of cognitive process, the idea of complementarity seems to point to. If it is true that the interactive nature of extended cognition is stressed out by second wave extended mind theorists in particular, by means of the "manipulation thesis" (Menary 2007, chapter 4), it also true that in Clark's work the idea of manipulation and that of the *active*

[33] For the discussion about the label "Extended Mind", see also Sutton 2014: 14.
[34] For my discussion of the locational claim see §I.6.

engagement with the external artifactual and social world is recurrent. For instance, Clark offers examples such as: the use of pen and paper to perform a long multiplication, the external manipulation of letters to play Scrabble (Clark 2008: 221), mapping the environment with tags to work on conceptual relations between objects (Clark 2008: 46), he uses the term "interaction" to describe Otto's cognitive engagement with his notebook (Clark 2008: xi), and so on. The problem with these examples is that, even if they are built on the assumption that the mind, at certain degree, unfolds through interaction (namely it has a relational nature), the functionalist explanation undermines this reasonable attempt. Indeed, the mind is said to be external because the little squares of scrabble, the paper with passages of the multiplication, and Otto's notebook function as an *external storage of information*. What is not clear is how "being stored" can be considered to be i) an activity, ii) a process; iii) something that points towards a relational engagement. In other words, the endorsement of a functionalist perspective undermines the epistemic effects of active externalism: to what extent externalities can be considered to be active if they count as cognitive, and then as mental, because they store relevant information for the deployment of cognitive processes? Even if their functional and causal role is as active as that of internal resources, passivity is still there in the background.

In conclusion, I think that the EMH is less radical than what has been thought in the development of the "4Es" debate. If we want to account for mind extend*ing,* other approaches to cognition, such as the enactive one, should be explored.

I.10 Concluding remarks. Convergences and tensions in the "4Es Cognition" debate.

In this chapter, I tried to draw a *fil rouge* that links different theories of the "4Es Cognition" debate. To do that, I offered a theoretical narration, aimed at sketching out the web of problems the debate unfolds by. That is why I began my discussion considering the way those theories react to classical approaches to cognitive science. Moreover, even if Cognitive Linguistics and Grounded Cognition are not "officially" part of the current debate I took into account, I considered those theoretical stances as well, in order to account for the broad framework Embodied and Enacted approaches to cognition come from. This choice is consistent with the general aim of this chapter: trying to give a broad picture of the debate, based on the historical and theoretical development of different theories, and on the consideration of the general problems different theories deal with.

This attempt of systematization was motivated by the way the debate unfolds. It is not uncommon to see Embedded, Embodied, Enacted and Extended cognition discussed together, as a whole. Even if, at a methodological level, understanding the theoretical links between the different approaches is fundamental, because this offers the conceptual tools to understand the debate and to actively take part in it, thinking that "4Es" actually are a robust philosophical coalition can be problematic.

My discussion of the core points of each theory was aimed at making this point clear. Certainly, there is an actual connection of problems within the debate. For example, the symbol grounding problem founds the Embodied/Grounded approach to cognition; the relevance of contextual factors is at the core of Embedded, Enacted and Extended approaches to cognition; the role of action in cognition is taken into account by all the theories of the debate; all the approaches deal with the problem of mental representation, reacting to classical cognitive sciences in different ways, and so on. Nonetheless, the ways each theory deals with the philosophical issues taken into account are very different. This draws different pictures of the mind, which sometimes integrate and reassess aspect of the other theories, other times offer explanations that are at odds the one with the other. Therefore it can be said that the voices of the debate are not harmonious, they do not sound like a *symphone*, a concord of sounds played by an orchestra. The "4Es" Cognition debate looks more like a jam session, of which improvisation, anticipation, mistakes, corrections, the uncertainty about whether a player can successfully join the other musicians, are core features.

The outline at the end of this chapter shows the intersections and the points of friction of this philosophical jam session.

Taking into account the main ideas gathered in this outline, it can be said that moments of tension can be individuated by considering (at least) three points: the role of the body in cognitive processes, that of representation, and that of action.

If the role played by the body in cognitive processes is taken into account, two strands of thought that entertain a relation of tension can be individuated (Clark, Kiverstein 2009:1-2). One of those strands focuses on the body at any level of its arguments. The body is considered to be the *very core* of cognitive processes. By focusing on the fined-grained details of the body, the full thesis of the Embodied Mind -which also the Enactive approach to cognition endorses- individuates the source of all cognitive phenomena in the materiality of the body and on the possibilities of action, inter-action, conceptualization, simulation and other cognitive activities that it produces.

Other theories, such as Embedded Cognition and the Extended Mind Hypothesis, have a different perspective on cognition. By resting on functionalist premises, they give no special role to the material substrates information is embedded in. According to EMH theorists, this functionalist perspective is precisely what guarantees the extension of the mind beyond the boundaries of the individual subject. According to supporters of this extended functionalist frame, this goal cannot be reached by the enactive approach to cognition, especially in its strong sensorimotor version (e.g. O' Regan, Noë 2001; Noë 2004). Indeed, according to the EMH, those approaches are too tied to the fined-grained details of the body (Clark 2008:177-179), and this makes them too "personalistic", in a certain sense. In order to have the same cognitive experience, two different subjects might have the same body, because this guarantees a kind of similarity in managing sensorimotor laws. But, if cognitive experience is so tied to fine-grained bodily details, how is it possible to think of cognitive experiences that are intersubjectively or materially extended?

If this idea suggested by the EMH is endorsed, it is seems possible to draw a "story of radicality" in the debate. This would start with Embodied approaches to the mind and it would end with the Extended Mind, passing by Enactivism. As said in the previous paragraph, I do not think that this story is true. Not only the EMH has some problems with mind extension, but also the

considerations about enactivism that EMH offers seem to be partial[35]. As explained in §I.7, also the enactive approach to cognition takes into account the problem of "mind-extending", and it does that in a way that, at least on a provisional level, seems to avoid the problems the EMH entails. By endorsing a complete relational account of cognitive experience, enactivism gets rid of the locational claim and of the introduction of passivity in the explanation.

These considerations bring me to the second point I would like to take into account: action.

In the article *Where's the action? The pragmatic turn in cognitive science* (Engel, Maye, Kurthen, Köning 2013), Engel, Maye, Kurthen and Köning claim that "4 Es" denote a "pragmatic turn" in cognitive sciences. Being action-oriented is a common feature all the theories taken into account share. This means

i) understanding cognition as the capacity of generating structure by action;
ii) looking at the cognitive agent as immersed in her task-domain;
iii) considering meaning-acquisition as tied to the context of action (Engel, Maye, Kurthen, Köning 2013: 203).

Now, all the theories examined in this chapter endorse these points. Nevertheless, by considering the issue of action, two strands of thought can be individuated again. Even if the majority of theories is aware of the fundamental role of motor action in cognitive processes and of the centrality of action in the context of cognitive practices, the enactivist approach to cognition builds its explanation on the notions of enaction and sense-sense making, namely on inter-action. As explained in §I.7, the focus on these conceptions of action is aimed to a theoretical substitution of internal representations with the relation between organism and environment. According to enactivism, the only viable

[35] It is worth to make clear that I think that Clark's way of considering the Enactive approach to cognition in his book *Supersizing the Mind* is partial and built "ad hoc" because he considers sensorimotor Enactivism only (in particular Noë 2004), and he makes use of his objections against the idea of full embodiment to introduce, and then to develop, his idea of Extended Mind. According to him, relying on functionalism, his explanation is said to be suitable to account for the continuous exchange of information between internal and external resources in cognitive processes (Clark 2008a: 219). Moreover, and this is not a critique to Clark, I think that this explanation of Enactivism, in retrospect, should be considered to be partial because Noë himself, which was Clark's target in chapter 8, recently extended his notion of skill (which, in *Action in Perception* was mostly tied to sensorimotor knowledge), by considering cognition as a skillful practice that is not only perceptual but that is also cultural and intersubjective (e.g. Alva Noë's last book *Strange Tools,* New York, Hill and Wang, 2015).

concept of action is external action, where by "external action" I refer to the externality of action in respect to the boundaries of the head, and not to something external to the organism-environment coupled system. Indeed, as already explained, the motto of the enactivist approach to cognition is "neither externalism, nor internalism" (Thompson, Stapleton 2009).

Theories such as Embedded Cognition and the Extended Mind, also focus on external actions, but, by endorsing extended computationalism and extended functionalism, describe them in a "discretized" way, as pointed out in §I.8 and §I.9. As I previously claimed, this undermines the project of giving a relational and dynamic account of cognition.

Moreover, the EMH seems to implicitly reassess some ideas about action developed in the theoretical frame of Grounded Cognition, and it does that when it talks about representations. As explained in §I.5, some versions of Grounded Cognition, such Barsalou's perceptual symbol systems, conceive of mental representations as modal representations described as internal re-enactments of sensorimotor activities.

The way Clark conceives of internal representations has something in common with this approach. As sketched out in §I.9, the idea of complementarity between internal and external resources leads Clark to think of Action-Oriented Representations (AORs). Those ones, as Barsalou's perceptual symbols, are representations strictly tied to action. They are defined as "*local* and *action-oriented* rather than objective and action-independent", as representations that "reflect the profound role of bodily motion [...] in shaping and simplifying the information-processing problems to be solved" (Clark 1997a: 149). As Barsalou's PSS, they internally reflect action and they are profoundly context-sensitive. According to Clark, this makes them interesting factors for the distribution of action (at least conceived in terms of function) between the internal realm and the external one.

A detailed consideration of this point will be offered in Chapter II. Indeed, that of representation is a huge philosophical problem. Moreover, it is particularly relevant for the "4E Cognition" debate and, as shown in the outline, it is a central point of tension between the Enactive and the Extended approaches to cognition, namely the two theories which explicitly point to "mind-extension". That is why, for the purposes of this research, this point should be taken into account in depth.

	Cognitive Linguistics	**Grounded Cognition**	**Biological Embodiment**	**Enactive Mind**	**Embedded Cognition**	**Extended Mind**
Body	Motor Body Body as a constraint	Sensorimotor Body Neurally represented Re-enacted Body	Material Body	Material Body Sensorimotor Body	Manipulatory Body	Functional Body
Action	Action as metaphor shaper	Neurally re-presented action Simulation Re-enactment	Motor Action Biological function	Enaction Sense-making Inter-action	Epistemic Action Action-Perception loops Action as computation	Epistemic Action Action-Perception loops
Context	Linguistic Situation Experienced context of linguistic practices	Re-enacted Context	No particular stress on this element	Transactions organism-environment	Manipulatory and symbolic context	Artifactual/ Technological context
Representation	Weak – Hybrid Representation	PSS B-Formats	De-Emphasis on Internal Representation	No representation	Internal + External Representation	Action-Oriented Representation External representation

Figure 3: Body, Action, Context, and Representation in second-wave Cognitive Sciences.

Chapter II

Activating Representations. Looking for the whole story

II.1 Introduction. Approaching Action-Oriented Representations.

In this chapter, I deal with the problem of representations between the Extended Mind Hypothesis and the enactive approach to cognition. In particular, I consider the problem of the relation between representation and action, which is one of the heated issues of the debate. Indeed, the Extended Mind Hypothesis claims that philosophical explanations need to account for action-perception loops by making use of the concept of action-oriented representation (AOR). On the contrary, theories such as "Radical Embodiment" and the enactive approach to cognition hold that AORs are not useful epistemic posits to solve the problem of selective responses to stimuli. This problem, also called "Generalized Relevance Problem" or "Frame problem" is the main reason why AORs have been introduced in philosophical explanations.

The ultimate aim of this discussion is to consider whether the way the EMH talks about AORs contributes to the philosophical project of "mind-extension" or if it rather prevents the explanation to come to radical conclusions about the nature of cognition and the mind.

In paragraph II.2, I consider the problem of practical knowledge in the "4 Es" cognition debate, explaining in detail the reasons that lead philosophers to talk about AORs, giving a representational account of practical-perceptual knowledge.

In paragraph II.3, I take into account Clark's version of action-oriented representation. I also consider the objections to non-representational accounts of action-perception loops. Those objections persuade Clark to endorse AORs. I suggest that Clark's arguments, in particular his notion of "representation-hunger", hide an unwarranted rhetorical slip.

Considered my worries about Clark's AORs, in paragraph II.4 and in paragraph II.5, I focus on the very concept of AOR in detail, in order to understand if it meets the criteria to determine if an item or a process is genuinely representational. Following Gallagher's article *Are minimal representations still representations?* (Gallagher 2008a), I claim that if AORs do not meet all these criteria they should not be considered to be representations. In that case the explanation should not buy AORs.

To develop my argument, in paragraph II.4, I consider the representational criteria endorsed by cognitive scientists, and I reassess them in a semiotic framework grounded in Peirce's thought. The outcome of this exam is a reduction of representational criteria from six to four.

In paragraph II.5, I apply this new set of representational conditions to Clark's AORs, and I show that they do not match three of these criteria. This means that they are not genuine representations; therefore a representational

gloss -as that one Clark makes use of- adds nothing interesting to our philosophical explanations.

Moreover, I consider the problem of AORs at a more general level of explanation, speculating about the role they play in the project of mind-extension. To the formal objections built around the idea of representational conditions, I add some considerations on the active features AORs are said to be endowed with. I show that Clark claims that he wants to make his representations active and dynamical, in order to make them fit with mind-extension, but he does not actually succeed in this project. Not only Clark's concept of action-oriented representation is tied to an internalistic prejudice, but also AORs are not active and dynamical at all. In the conclusion of my discussion, I suggest that a possible solution to this lack of dynamicity and activity in representation can be found in Peirce's concept of sign. This discussion of Peirce's concept of sign is meant to be a preparatory stage for what will be considered in the following paragraph.

Indeed, given the problems that Clark's AORs entail, in paragraph II.6, I consider a possible solution to the generalized relevance problem, which is a philosophical proposal that does not make use of cognitive scientists' AORs. This account of selective responsiveness to the environment crosses some Merleaupontian insights on bodily intentionality with ecological psychology, and explains cognitive agents' ability to act meaningfully in a context by focusing on the concept of affordance. I claim that this is an interesting approach to the problem of relevance, and that it does not entail many of the problems Clark's explanation suffers from. Nevertheless, I suggest that this "radical embodied" or enactivist explanation can be made stronger and more complete if it is integrated with Peirce's concept of indexicality.

In §II.7, I offer some concluding remarks concerning the notion of AOR. Here I claim that not only the enactivist approach to contextual relevance (especially if it is integrated within a semiotic framework) is the more viable one, but it is also suitable to account for the mind's extension.

II.2 Acting without representing?[36] Practical knowledge and minimal representationalism.

In Chapter I, I discussed the main positions of the "4Es Cognition" debate, explaining why, even if those theories share a common theoretical ground (i.e. their critical reactions to internalism and methodological individualism defended by first-wave cognitive sciences), they cannot be considered to be a robust philosophical coalition.

In particular, I considered three main points of friction, which constitute great factors of differentiation: the role given to the body in cognitive processes, the relationship between action and cognition, and the use of the concept of representation.

I also noticed that both the enactive approach to cognition and the Extended Mind Hypothesis are theories which, relying on different explanatory strategies, endorse an "extensive"[37] explanation of the mind. When I say that both endorse an "extensive" view of the mind I mean that their claim is the following one. In order to understand what the mind is, we should explain how cognitive activities unfold by considering the cognitive role of personal resources (i.e. somatic and neural resources) and extra-personal resources (i.e. extra-somatic resources, such as information distributed in the environment or embedded in cultural and technological artifacts, in Clark's explanation, and the space of intertwinement between the agent and the environment, according to the enactive perspective).

Nevertheless, a closer look at those explanations aimed at accounting for the process of mind-extension shows that these two theories are radically different. In particular, what seems to make the real difference is the way these two approaches deal with the notion of *representation* in relation to that of *action* or *activity*.

My working hypothesis about this point is the following one. In order to understand if and how those philosophical theories actually account for the mind as an extensive process, we should look at the way they deal with mental representations. Indeed, the internalistic stance of mainstream cognitive

[36] The title of this paragraph recalls Clark and Toribio's article "Doing without representing?", *Synthese*, 101, pp. 401-431, 1994.

[37] Here I borrow the expression "extensive" from the article "Extensive enactivism: Why keep it all in?" (Hutto, Kirchhoff, Myin 2014) and from Hutto and Myin's book *Radicalizing enactivism* (Hutto, Myin 2013, Chapter 7) in order to avoid the problems the expression "extended" entails. Those problems have been already discussed in Chapter I, §9. This is my personal terminological choice. Indeed, some scholars which argue in favor of cognition/mind-extension from an enactivist perspective or from a Radical Embodied view of cognition, and not from a functionalist one, seem not considering this kind of problem (e.g. Chemero 2009: 31-32; Colombetti 2015).

sciences depends on two philosophical claims built on the concept of representation. The classical cognitivist-fashioned argumentation is this one.

i) Representations distinguish the mind from the rest of the world. A state can be considered to be a mental state (namely a cognitive state, conceived as a propositional attitude that causes behavior) when it is endowed with a content that represents what is outside of the mind, namely when it establishes a mediation between the cognitive agent and the environment.

ii) Cognitive representations set the boundaries of the mind within the limits of the brain. Representations can be defined as mental if and only if they vehicle naturalistically determined (namely non-derived) contents (Menary 2010c: 15). If a process does not involve intrinsic content, then it cannot be considered to be cognitive (Adams, Aizawa 2010b: 70). Indeed, representations that bear contents in virtue of norms socially established (think about road signs, written language, tags, and so on) have a representational function that is derivative only: their meaning-bearing feature is "parasitic" (Clark 2010b: 48 for this expression), not original. Their "standing-for" is a feature that is ascribed to them from the outside. Then, according to those philosophical views, to talk about cognition properly, we should seek for items whose representational function is original, underived. Those representations are identified with neural states (Adams, Aizawa 2001; 2009; 2010a; 2010b): they are biological resources that *naturally* stand for the inputs that activated them. Developing their argumentation in this way, classical explanations of the mind, and also more recent brain-bounded approaches to cognition, exclude any possibility for mind-extension.

Now, as sketched out in the previous chapter, the EMH brings into question the very idea that intrinsic content is a necessary condition to talk about cognitive processes properly and, at many explanatory levels, it rejects the idea that neural representations are the very core elements of cognitive processes. Nevertheless, it claims that there are some cases in which cognitive systems strongly rely on internal representations (e.g. when the cognitive system reasons about the counterfactual and the abstract; when the cognitive system deals with an informationally noisy context). Those kinds of internal representations are reworked in order to meet the core ideas of the EMH: they are said to be highly context-sensitive and action-oriented. Despite this reassessment of mental representations, I think that this philosophical move does not actually hit the

spot: the way the EMH thinks of representations still entails some of the problems classical notions of representation suffered from (e.g. it still accounts for cognition as the ownership of mental states and not as an activity). This is a flaw in the argumentative structure of the EMH, and it undermines the general project of "mind-extending".

Before considering the problem of "active representations" in the "4Es" literature in detail, taking into account the EMH proposal and the objections against this view made by enactive approaches to cognition, it is worth to offer a short introduction to the philosophical ground the idea of active representation comes from. This is useful to understand why some theorists think that we still need a minimal form of representationalism. Therefore, I will devote this paragraph to a general discussion of the enactivist anti-representationalist stance and to consider the reasons that persuade some of the voices of the "4E cognition" debate to re-introduce mental representations as epistemic posits.

As already explained, the enactive approach to cognition describes the mind as a relational process through which, in virtue of the transactions between the organism and the environment, sense (namely "living meaning", a direction for action) is created. In other words, the mind is explained as generating structure (namely what I explained in my discussion of the problem of normativity in relation to the concept of Umwelt) through action. The mind is made of the continuity of the organism's actions in an environment that, by means of those interactions, becomes part of a coupled, structured and structuring system, made out of the human cognitive agent and the space of salience of her activities. The mind is then understood as a process of creation (Varela, Thompson, Rosch 1991: 148), as the enactment of meaningful relations, which cognitively define both the agent and the world.

This explanation radically differs from what Varela, Thompson and Rosch define "cognitive realism". Cognitive realism claims that "the world can be divided into regions of discrete elements and tasks" (Varela, Thompson, Rosch 1991: 147), namely into sections that have certain given (functional and computational) properties.[38] Mind and world, according to cognitive realism (which the authors of *The Embodied Mind* identify with approaches to cognition such as cognitivism), should be described as independent the one from the other (Zahidi 2014: 461). There is a given world, endowed with its own properties, and a mind, different from the world in its materiality, features, properties and functions. Given this independence of mind and world, the

[38] It is worth to notice that, following what I said in §I.9, the EMH in Clark's version fits with this definition of cognitive realism, and then it can be criticized relying on this enactivist argument.

mind, in order to acquire a cognitive grip on the world in which it acts, is said to internally represent properties, features and aspects of the world. By means of mental representations, human agents re-present, that is make present again, what is already present in the external world.

As noticed in the discussion offered in §I.7, the concept of enaction is a critical reaction to this picture of the mind entailed by cognitive realism. My discussion of the role played by the context in cognitive processes, in which I metaphorically recalled the idea of the relation between figure and background according to Gestalt psychology, was precisely aimed at making this point clear. Conceiving of the space of cognitive practices not as an objective space endowed with pre-given properties, but rather as an Umwelt, the enactive approach to cognition gives a justification to the choice of cleaning up its explanatory framework from representations. The Umwelt is not a world that has pre-definite boundaries, but it is rather a *negotiable space* in which meanings vary according to the possibilities of action of an embodied agent, who is defined by the historicity of her personal situated experience. That is why cognition is said not being dependent on internal representations that mirror an external structure. The world where cognitive practices take place does not actually "exist" before the enactment of those cognitive practices.

Then, according to Enactivism, it follows that there is nothing to re-present, namely to make present again, but there is a lot to create. As I explained at length in §I.7, there is a coupled system that has to be organized, namely brought forth. There is a space of meanings, of salience, invitations for action, which has to be created. And also, and *crucially*, there is a basic and fundamental "knowledge how", an implicit "readiness to hand" (Varela, Thompson, Rosch 1991: 148) that has to be first achieved, and then enriched through experience. To explain this last point better, I would say that, in order to make the continuous *looping*[39] between the organism and the environment possible, namely to guarantee the continuity of life and cognition, not only the environment has to be structured, organized in a certain way, viz. in a way suitable to satisfy the organism's needs, but also the organism itself has to acquire a set of *habit-based skills* able to actualize the structural attunement between the agent and the environment. This is to say, through experience (where "experience" corresponds to "practice"), not only the world is structured in a certain way (e.g. as a habitat for the development of the organism's life; as a familiar space, disseminated of the products of a culture)

[39] "Looping" is an expression that is more common in the literature about the Extended Mind (e.g. Clark 2008; Clark 2001; Clark 1997; Clark 1998) than in the enactive literature. Nevertheless, recently it has been used also by philosophers that endorse a critical stance towards the EMH (e.g. Noë 2015). To me, "looping" is a good word to refer to the intertwinement of the agent and the environment, because it accounts for its processual nature, it points to the continuity of these processes, and it points also to the realm of activities.

but also the human agent is changed, modified or organized in a certain way. Experience makes the agent able to support the dynamical balance with the contingent fluctuations of the environment (Cappuccio, Froese 2014: 5), endowing her with a fundamental kind of practical knowledge that makes her actions attuned to, and then relevant in an Umwelt. This one is not conceived as a physical space only, as a perceptual space or as a motor space, but it is also social and cultural. It is an Umwelt that people share in their being human and in their being shaped by culture (Varela, Thompson, Rosch 1991: 172-173).

Here it is very important to pay attention to the term "attunement". As already sketched out in §I.7, when the enactive approach to cognition explains how human agents have or acquire a grip on their reality, sometimes it describes this process in terms of attunement (e.g. Bruineberg, Rietveld 2014; Cappuccio, Wheeler 2012; Degenaar, O'Regan 2015; Hutto 2006; Gallagher 2017; Ward, Stapleton 2012 for the use of this expression). Enactivism does not appeal to the idea of correspondence between the subject's internal states and states of affairs or events of the world, correspondence whose "efficacy conditions" depend on an internal representational content which can or cannot fit with what is out there, in the external world.

It is interesting to notice that "attunement" is not a new expression in the history of philosophy. Looking at its use in another philosophical approach can be useful to understand what "attunement" means in the enactive framework.

"Attunement" is a translation of the word *Befindlichkeit*[40]. *Befindlichkeit,* according to the terminology of Heidegger's *Being and Time,*[41] is the fundamental "characteristic of finding oneself in a world through a mood" (Ratcliffe 2013: 157). Attunement is a fundamental *existentiale*. It pertains to the ontological constitution of *Dasein*. This means that it is not conceived as something that is experienced or owned as a state of mind by a psychological

[40] Other ways to translate *Befindlichkeit* in English are "disposedness", "sofindingness", "affectedness" (see Ratcliffe 2013 for a more detailed discussion of this point).

[41] At first glance it could seem to be the case that this introduction of the term "attunement" in relation to Heidegger's work is not consistent with the discussion I am offering here, which focuses on Philosophy of Mind and on Philosophy of Cognitive Sciences. Nevertheless, I think that this consideration is misleading. This for three reasons. First, enactivism is strongly influenced by the phenomenological tradition, of which Heidegger is a leading figure. This, at least at a general level of analysis, justifies the idea of considering some Heideggerian concepts to understand the philosophical roots of the problem I am dealing with. Moreover, although many works that endorse an enactivist perspective consider Merlau-Ponty's thought more than Heidegger's one, recent works in philosophy of cognitive science take Heidegger's work as an important source for the development of their research (e.g. Malpas 2000; Kiverstein, Wheeler 2012). Third, the issue I am considering in this chapter (namely new ways of conceiving of mental representations) has its roots in a discussion about non-representational artificial intelligence developed in Heideggerian terms (Dreyfus 2002; Dreyfus 2008; Cappuccio, Wheeler 2012; Wheeler 2005).

subject, nor it is experienced "out there", in the world. Contrarily, attunement means "being part of a world that is pre-subjective and pre-objective, [and] all 'states of mind' and all cognitions of 'external' things presuppose this background sense of belonging to a world" (Ratcliffe 2013: 157-158; text into brackets added). To put it in other words, "attunement" is what lays the foundations of our possibility of being engaged (viz. knowing how to cope) with the different practical meanings that worldly situations offer. Indeed, "attunement" refers to the ontological structure thanks to which meanings can be appreciated as "inextricable from our actual and potential activities" (Ratcliffe 2013: 158). "Being attuned" means finding ourselves as sensitive, responsive, ready to act upon the world in a certain, meaningful, way. This ability that *Befindlichkeit* founds is not a kind of propositional, intellectual-like, or "staring-like" knowledge about the world, but it is rather pre-theoretical, and it is always and already practical.

There is a passage from *Being and Time* that could be useful to explain this point better. Heidegger writes:

> "the less we just stare at the hammer-Thing, and the more we seize hold of it and use it, the more primordial does our relationship to it become, and the more unveiledly is it encountered as that which it is—as equipment" (Heidegger 1927/1962: 98)

Heidegger claims that our primordial relationship with the world, namely the most important way the encounter between the agent and the world unfolds by, takes place in practice. In the average everydayness of experience, that is what "4Es" approaches to cognitive sciences would define as ordinary cognitive experience (and, as explained in §I.2, this is what second-wave cognitive science aims at studying), we do not passively stare at what is in front of us. We do not encounter discrete objects, things that we should intellectually discover, by objectifying their features, namely by describing how they look, *saying that* they serve this or that purpose, and so on. On the contrary, in virtue of the constitutive ontological dimension of attunement, we encounter the world as ready-to-hand or as being-at-hand (*Zuhanden*), as actively graspable, and then as always and already cognitively graspable through action, and, in the case of the hammer, through manipulative action.

Explaining what Heidegger says in the passage cited above, I would say that when we encounter a hammer, instead of looking at it and asking ourselves "what is it for?", and answering to that question by saying "it's just a hammer"[42], ordinarily, we just *use* it, we inter-act, we pre-reflectively act-*with* the hammer

[42] This passage is adapted from Richard Schmitt's article "Heidegger's Analysis of 'Tool'" (Schmitt 1965: 73).

because we are disposed (attuned) to react to the "hammering of the hammer" (Schmitt 1965: 75). This is to say that we encounter the world as a rich field of '*in-order-to's*. We encounter objects that point to their manipulability, they display their goal-directedness, and, in doing so, they point to the significance of the possible world-engaging practices our experience unfolds by.

We are open to, or, to use Heidegger's terminology, we are "thrown in a world" of possible practical meanings. By the phrase "practical meaning" I refer to the possibility of making sense of the world through actions, where the condition of possibility of those actions has to be found in our *circumspective* way (*Umsicht*) to look (*sicht* means "sight") at the world, or better, of being engaged with the world.

Now, what has been said about the hammer can be somehow extended to all objects in humans' practices. Obviously, the hammer is a peculiar case study, because its being a tool makes its constitutive feature of "pointing to" to the web of human practices more clear. Nevertheless, what Heidegger says about attunement, circumspection and practice seems to make sense for any kind of encounter with the world. Indeed, according to a Heideggerian perspective, the encounter with the phenomenological reality is said to be tuned by a kind of *mattering* (Ratcliffe 2013: 159). Things show up as "this" or "that", they stand out as distinct the one from other, because they matter for this or that kind of practice. They are relevant to the context because they matter, and their appearing as mattering is dependent on that *Umsicht*, on that practical and primordial look through which the world and the web that connects things in the world is lived and acted.

To explain this point in another way, it can be said that the world shows up as meaningful and as something that matters through action (or through the possibility of action). Actions are performed in a way that has a relevance that pertains the context of practice because there is a practical, pre-reflective, non-intellectual background knowledge that makes us sensitive to contextual relevance.

Now, a similar approach to the issue of practical knowledge and to the way this makes the cognitive agent sensitive to contexts, namely it makes her ready to act meaningfully in a given context (then "actualizing a condition of attunement"), can be found in some enactivist approaches to cognition.

As said previously, it is common to find the expression "attunement" in the enactivist literature. Even if enactivist scholars do not always endorse an existential perspective as the Heideggerian one and they do not explain "attunement" relying on an ontological dimension, their way of thinking of this idea recalls some of the points that I briefly sketched out here.

According to the enactive approach to cognition, thanks to the repeated interactions of the agent with her Umwelt, the agent acquires a set of *skills* that she habitually displays during her practices. This makes the practical attunement between the agent and the environment continuous, spontaneous, taken for granted, unreflective in practice.

To the enactive approach to cognition, the concept of skill, which seems to play the same epistemic role of Heidegger's concept of *circumspection*, is firstly understood in its fundamental relation to the body: skills are embodied.

This embodied conception of skill is already sketched out in *The Embodied Mind* by Varela, Thompson and Rosch. The authors of the book talk about *motor skills* (Varela, Thompson, Rosch 1991: 147) as the fundamental embodied ability that makes agents able to cope with a changing environment, foreshadowing the huge debate about sensorimotor approaches to perception that is now one of the heated subjects of second-wave cognitive sciences' debate.

Here is my way to unpack their argument. If it is true that, in some particular cases, cognitive experiences (conceived as problem-solving tasks) take place in a space with defined limits (think about the defined space of chess in chess playing tasks, for example), our experience usually develops in a less circumscribed domain, whose parts seem to be not defined at all. Moreover – and this is a crucial point- the environmental space of everyday cognitive practices is characterized by a high degree of mutability and unpredictability. Given this complexity of the environment where cognitive tasks take place, the authors of the book claim that it is unlikely that cognitive agents acquire a grip on the reality in which they act by mentally representing it in a detailed way. Indeed, the space of the cognitive task is not discrete and its details change quickly. Then it is not clear to what extent mental images (representational items with defined boundaries) that correspond to the environment or to parts of it can be produced efficiently.

Moreover, given the dynamical and changing nature of that space, if we try to explain how the agent grasps her world by representing it, we are forced to suppose a really complex system of internal representations, which changes in tandem with the worldly domain. This is not epistemically cheap, because this kind of explanation requires the agent to be endowed with an extremely sophisticated system of internal representations, whose nature is not actually clear. How does the individual sensory impressions are connected the one with the others? How much work should the mind do in order to produce a synthetic image of the perceived space? To what extent and in which way this image corresponds to external reality? In which way this static image of the external reality affords action in real-time cognition?

Then, it is more intuitive to think of the possibility of coping with the changing environment relying on the notion of practical knowledge or knowing-how, for example describing the activity of successfully moving in a space of ever-receding levels of detail by directing bodily movements and by attuning them to contingent circumstances in virtue of motor abilities, which structure the perceptual space in order to satisfy the cognitive agent's expectations. For example, the agent moves her head, eyes or her whole body to have an access to what is present in the world, performing a motor task in a successful way.

This idea according to which the agent's *perceptual attunement* with the environment is dependent on an implicit knowledge that habitually structures the environment is then developed widely in more recent enactivist-fashioned approaches to cognition (Noë 2004; Noë 2010; O'Regan, Noë 2001). For instance, Alva Noë, in his book *Action in Perception,* develops a non-representational account of perception[43] based on the concept of skill. The notion of skill is understood primarily as *bodily skill,* and it is strictly tied to the notion of *sensorimotor knowledge* (Noë 2004: 12). Indeed, the core idea of this approach consists in claiming that, in our perceptual experience, we do not represent contents (such as shapes and spatial relations) but we rather *enact* them, and we do so thanks to the mastery of sensorimotor knowledge. Sensorimotor knowledge is the practical embodied knowledge of the way the spatial relations between the cognitive agent's body and objects in the environment vary according to bodily movements. By moving our body in a certain way, we enact "contents" that are virtually present in the environment.

According to this philosophical stance, perception is like the activity of painting, namely a skill-based, world-engaging process of exploration in which the eye probes the scene, then it goes back to the painting, and then again to the scene, and so on. This comparison between perceptual experience and the activity of painting is aimed at explaining that perception takes place in a circle in which *we do what we perceive*, and we do that in a "good way" because we "*implicitly know*"[44] the bodies we have. We know how to move around in order

[43] Here it is important to notice that, by recalling the phenomenological tradition, he claims that perception founds any kind of cognitive experience, included what we do with concepts (Noë 2004, Chapter 6).

[44] It is worth to point out that, at first glance, one could be tempted to interpret this kind of knowledge by recalling the concept of *image schema*, notion I took into account in §I.4. This interpretation of the agent's knowledge of her body, in the case I am taking into account here, is misleading. Indeed, the kind of knowledge I am referring to does not entail any residual reference to representational contents. Those ones were implied in image-schemas, which were defined as "dynamic analog representations of spatial relations and movements in space derived from perceptual and motor processes" that are intentional objects of consciousness (Gibbs, Colston 1995: 349). In contrast to image-schemas, the kind of knowledge Noë talks about is non-representational, and it seems that it can be described well by making use of the

to make present what is given as virtual in the environment because we "possess the sensorimotor knowledge to be *effective* in our exploration" (Noë 2004: 75, italics added).

An example that Noë offers is the following one (Noë 2004: 77 and followings). Consider an opaque, solid object, like a tomato. When this object is seen, it has visible and invisible parts. The agent, since she is always situated in the environment in a certain way, namely according to her bodily constitution and posture, can see part of the surface of the object only (for example its facing side). Now, if we try to explain the perceptual (visual) experience of the tomato by considering sensory impressions only (maybe conceived as metal representations of the sensory data present in the world), we would probably find ourselves in a theoretical situation that sounds a bit paradoxical. If we report our perceptual experience of the tomato, we can talk about an object that is round, voluminous, three-dimensional. How is this possible? If perception was a result of an internal association of representations of sensory data, the image of the external object that we would represent would be bi-dimensional. Bu this is not the case. It seems more luckily that we experience parts of the tomato that, strictly speaking, we do not see, and then cannot be manipulated by the process of "internal translation" that starts with sensory data and ends up with mental images or representations.

According to Noë, visual experience has this structure (viz. we perceive solid objects as solid even if we see them just from one perspective; we perceive something that makes sense to our experience even if the bottom-up data alone seem to suggest another story) because perception is supported (and made possible) by sensorimotor skills. We know how to move our bodies in order to make present what is given in a virtual modality in the environment. We implicitly know that, if we turn our head a bit, the spatial relation between us and the perceived object will change. We implicitly know that we will have access to another side of it.

To unpack this point, Noë writes:

---

concept of *body-schema*. By "body schema" I mean a "non-conscious performance of the body", a performance that is not an intentional object present to the subject's consciousness, but that is rather a performance by means of which the body acquires a certain kind of organization in relation to the environment. "The body schema is an active, operative performance of the body, rather than a copy, image, global model, or conception of existing parts of the body. [...] [It] is the body as it actively integrates its positions and responses in the environment" (Gallagher 1986: 542, 548; text into brackets added). Body schemas, rather than entailing a kind of intellectual understanding of one's own body, are the pre-reflective, non-conscious organization of our bodies in relation to the environment, and consequently in relation to our possibility of inter-action with the environment.

> "To experience the figure as a cube [or as a voluminous object, such as a tomato], on the basis of how it looks, is to understand *how* its look changes as you move" (Noë 2004: 77; text into brackets added)

Perception consists in an active process enabled by an implicit, non-conscious, ready-to-hand practical understanding of the way our actions modify how the environment looks like, and then our relation with parts of the context of the cognitive practice. Our perceptual experience unfolds smoothly, quickly, directly (and this seems to point to a relation of attunement between the perceiver and the environment in which the perceptual experience takes place) because we are skillful embodied agents that act upon, or act out (enact) the environment according to an embodied knowledge that is always present in the background of our experience.

This account of perceptual experience is widely accepted by most enactivist scholars, and, even if some points of this theory are not widely developed by each individual version of enactivism, or they are explained in different terms[45], what all the enactivist scholars consider to be at the core concept of their approach is the notion of skill.

Skill is a concept useful not only to account for perceptual experience, but it also lies at the core of explanations of other aspects of cognition, usually defined as higher-order cognition. For example, this idea that skills are at the very core of our cognitive experience has been developed further in the so-called enactive tradition to account for more complex kinds of experiences of attunement, such as collective practices.

For example, Gallagher and Ransom, in their article "Artifacting Minds: Material Engagement Theory and Joint Action", integrate some enactivist insights with Material Engagement Theory[46] (MET) to explain how collective action takes place. In this article, they endorse a critical stance towards theories of joint action that preserve intellectualistic or internalistic traces. They claim that, instead of trying to explain collective action by positing a set of sophisticated conditions[47], it is better to endorse a more embodied approach,

[45] See for instance the huge problem of the relation between cognition and consciousness, or the use of expressions such as "actual" and "virtual", very common in Noë's books but not in other enactivist approaches to cognition.

[46] See for example Malafouris, L. (2013), *How things shape the mind. A Theory of Material Engagement*, Cambridge (MA), The MIT Press.

[47] Such as individual plans or individual intensions that forerun shared beliefs that structure joint action; or the idea that individuals with separate intentions form plural subjects relying on normative commitments to act together; or the idea that collective agency takes place when there are individuals with interlocking intentions and meshing sub-plans (Gallagher, Ransom 2015: 341).

which also and crucially pays attention to the cultural dimension of skillful interaction.

To explain their argument, it is worth to examine one of the examples they offer (Gallagher, Ransom 2015: 345 and ff.). Think about the joint action of moving a bookcase. This is a kind of action that - given the bodies we have and the material structure of the artifact the action deals with- cannot be performed by individuals separately. It is an action whose efficacy heavily depends on the coordination of the two subjects' bodily movements. Moreover, the success of the joint action heavily depends on the material structure of the artifact the action deals with, and also by the material structure of the environment in which the practice takes place.

For example, to move the bookcase from one room to another, the participants of the joint action have to get the bookcase through the door. This entails an attunement of movements. I feel that the bookcase starts to go on the right, that is your left, and so I automatically change the position of my hands, or I change the amount of force to apply to that part of the bookshelf. We jointly deal with this practical task thanks to our sensorimotor skills. I react to your movements quickly because I implicitly know the possibilities of my body, and I feel the possibilities of yours thanks to the manipulation of the material artifact. I understand what you are doing (and this matters for the development of our joint practical task) because we are jointly manipulating the same object, which functions (at least) as a medium for the communication of spatial and motor relations.

Since collective action is not the main topic of this chapter, I will not explain the philosophical implications of this approach further. Nevertheless, it is worth to notice again how much the enactivist perspective broadly understood[48] emphasizes notion of skill (and consequently that of attunement, as I already explained). Also when it explains complex cognitive practices, such as collective action, it gets rid of internal representations, which, in the domain of study of joint action, could be understood in terms of plans mentally represented, as the internal representation of the other agent's intentions, or as the mental representations of features of the objects manipulated in the practical engagement. The emphasis is rather on skills the agents share. Primarily, sensorimotor skills, and then on what can be interpreted, to some extent, as "social skills".

Focusing on this last idea of skill they write: "In some regard, when I add my skills (strengths) to your different set of skills (strengths), we both may gain by expanding the affordance space to offer more possibilities" (Gallagher,

[48] I say "broadly understood" because in Gallagher and Ransom's article there's a lot of MET, theory which combines the enactive approach with distributed, extended cognition and cognitive archeology (Malafouris 2013: 3).

Ransom 2015: 347). It is in this sense that skills become "social". Through interaction, agents exhibit their skills, and, by showing them, they share them. In doing so, they create a new space of affordances, a space that solicits actions whose performance was limited or prevented before the joint action took place. Think again of the example of the bookcase. The relations between the agent's possibilities of action and the environment[49] -conceived as *situation*, namely not objectively (i.e. as a physical space that can be described by individuating its properties), but as the agent-context system- before the exhibition of joint skills are different from those emerging from skillful collective performances. Without you, the bookcase in the room points to a different field of affordances, namely to different kinds of possible actions, holistically connected. The situation taken into account can be described as disseminated of invitations for action such as: reading the books situated in the bookcase, appreciating the design of the bookcase and its harmonic relation with other pieces of furniture in the room, climbing on the bookcase shelves to grab the DVDs on the last shelf, and so on. There are many actions that can be performed in that situation, but the field of affordances made of the abilities of the individual agent and the environmental situation does not afford the action "moving the bookcase". On the contrary, it prevents this action from being performed. The skillful joint action changes and enriches this landscape of affordances (Rietveld, Kiverstein 2014, for the use of the expression "landscape of affordances"). The joint activation and coordination of motor skills makes actions such as turning the bookcase, lift it, slide it on the floor and so on, possible. To put it in other terms, the practical meaning of the object "bookcase" changes or is enriched by means of inter-action, because, relying on "collective abilities", the set of actions our relation with the bookcase affords changes (namely what can be done with the bookcase changes). A new space of meanings, and then for action, is created.

After this short explanation of the role of practical knowledge, skills and abilities in cognitive processes according to the enactive approach to cognition, some general considerations can be offered. The concepts of implicit practical knowledge and that of skill, considered in relation to different cognitive levels

[49] This is an interpretation of the concept of affordance offered by Antony Chemero (Chemero 2003; Chemero 2009) in order make this notion clearer. Indeed, in its original formulation developed in *The ecological approach to visual perception* by James Gibson (1979), the notion of affordance oscillated between being a source of the environment (Gibson 1979/2015: 127) and being the relational space between the agent and the environment (Gibson 1979/2015: 129). Other contemporary interpretations of the concept of affordance focus more on the "invitation-aspect" of parts of the world, which are considered as available resources for possible actions (Rietveld, Kiverstein 2014: 327, which cite Reed 1996 and Silva, Garganta, Araújo, Davids, Agular 2013).

(e.g. perception, collective action), seem to play a kind of "foundational" role in cognition. We know what we know, and we get to know the world in this way, because we are embedded in cognitive situations according to a certain a kind of *attunement*, which is always practical and active. Even perceptual experience, which along most part of the history of philosophy was said to entail a certain degree of passivity, is explained as an active process grounded on practical knowledge.

This focus on practice, which not only pertains human cognitive abilities, but also the way the environment shows up as practically loaded, is strictly connected to the issue of representation in cognition, as already sketched out taking about Varela's et al. (Varela et al. 1991) objections against cognitive realism. This stress on action in cognition is a reaction to a certain way of conceiving of cognitive processes as relying on *passive representations*, namely representations modeled as words, which bear meaning in virtue of their standing in appropriate relations to things extrinsic to them, thanks to an external action of interpretation performed by something- an interpreter- that is "outside" of the representation (Rowlands 2006a: 2). Moreover, this way to explain cognition and how the cognitive agent is attuned with her world, represents a critique to explanations of cognitive processes that account for the relation between agent and world by appealing to *internal mediations* more in general.

Now, as anticipated in Chapter I, this attention paid to practice is not an exclusive idea of the enactive approach to cognition. As the outline at the end of the chapter shows, the broad project of second-wave cognitive science points towards this direction. Nonetheless, it seems to me that there is one approach to cognition that tries to be as faithful to practice as Enactivism is: the Extended Mind Hypothesis.

For example, one of Clark's core ideas is that the mind extends into the world thanks to processes of looping, which are grounded in action. Some examples of looping are: what we *do* with pen and paper, performing a mental activity in which the "internal realm" and the "external" one reciprocally influence each other (Clark 2008: xxv); agents that improve their conceptual capacities by manipulating external media, such as material symbols or tags (Clark 2005a: 257); improving mathematical skills by manipulating external medias, such as an abacus, or more in general, mathematical notations (Clark 2015: 277).

The examples taken into account share some characteristic features.

First, they entail the concept of skill, directly or indirectly. Learning, thinking, conceptualizing are dependent on the acquisition of an *ability*, of a kind of knowing-how, that gives sense to practice. To use pen and paper to

think, the agent has to be able to manipulate pen and paper, and she has to know how to work with arbitrary notations (such as the letters of a language). The same things can be said about the example of tags, and about the example of mathematical notations.

Second, all the examples implicitly rely on the idea that skills can be acquired and improved because humans structure the environment in a certain way. In doing so, the environment becomes a space in which cognitive tasks can be performed smoothly, quickly and efficiently. Pen and paper, the abacus, symbolic notations modify the cognitive space of our tasks because they offer us a space to display our skills smoothly. Skilled agents and scaffolded environments constitute what has been called a coupled system, notion that is not very far from what I called "attunement".

Nevertheless, despite this attention to practical knowledge (conceived in terms of skills, active manipulation, action, sensorimotor knowledge and as non-verbalized knowledge; Clark 2008, Clark 2015; Ward, Roberts, Clark 2011), there is a fundamental element that distinguishes this perspective from the enactive one. Skillful action is said to be mediated by internal representations that, in contrast to classical notions of representation, which implied passivity at a certain degree, are said to be *active* and *action-oriented*.

The introduction of the concept of action-oriented representation not only depends on the pluralistic frame in which the EMH is developed and on the core ideas of this theory (e.g. cognitive impartiality and complementarity), but it also constitutes a critical response to the way the enactive/radical embodied approach to cognition explains how the cognitive agent and the environment are practically attuned.

The issue that critiques of fully anti-representational accounts of cognitive practices consider is the following one. Acknowledged that many of our cognitive experiences have a practical core, and acknowledged that our ability to cope with our environment depends heavily on the knowing-how that we embody, is skill in itself enough to explain how we act *meaningfully* in a given context? Skills are something that cognitive agents acquire through experience: skillful agents are endowed with sets of *specialized* knowing-how, which make sense in a particular ad defined context. There is a subtle difference between claiming that agents cope with their environments by means of *practical knowledge* and saying that they display *exactly* that kind of skill that makes their actions meaningful, relevant in that particular context (Wheeler 2010a: 333). How can we explain the fact that humans usually exhibit the *right kind* of skill the practical situation requires?[50]

[50] This is a formulation of what is called *Frame Problem*. Broadly speaking, the Frame Problem is the problem of *cognitive relevance*. There are many versions of the frame problem. When I talk about a cognitive agent that has to choose among different kinds of skills, I refer to the

To understand this point better, it is worth to notice that the kind of practical knowledge the word "skill" refers to seems to have two fundamental aspects. On the one hand, it is described as an *implicit* practical knowledge that the agents *embody.* It has a pre-theoretical nature: it becomes *habitual* through experience. That is why the agent does not need to get access to it by means of an intellectual act (e.g. by mentally representing the rules this kind of knowledge unfolds by; Dreyfus 2002b: 367). This kind of knowledge, trough time, becomes part of the agent's non-reflective, smooth acting in familiar environments (Rietveld 2008b). On the other hand, if it is true that this practical knowledge is something that agents embody, and in this sense it is something agents are endowed with (even if they do not usually speculate about it), it is also true that skills heavily depend on the cognitive situation in itself. Skill is a kind of *situated knowledge*, in the sense that it is not only triggered or activated by aspects of the environment, but it is also realized in the space *between* the acting agent and the environment. The word "skill" points to a kind of relational practical knowledge that is constitutively context-dependent, in the sense that occurrences of the habitual knowledge agents embody are solicited, activated and made meaningful by what is there, in the environment. In this sense, agents can be said to display the right kind of skill in that given situation because the practical knowledge I am talking about is relational. It is not just an ability agents are endowed with, but it is also something the environment affords. What is in the environment is part of the realization of this practical knowledge.

Think again about Uexküll's "performance qualities", taken into account in §I.7. The core idea of Uexküll's theory of meaning consisted in claiming that the world, conceived as an Umwelt, is not a neutral space, but it is a meaningful space endowed with active, practical qualities. The world shows up as active, or capable of being activated under certain respects (e.g. the throw-quality of the rock, the path-quality of the country road, and so on). In doing so, namely in its being active in *specific* ways (performance qualities), in its entanglement with the skillful agent, it realizes a specific kind of knowledge, relevant for that peculiar situation. To put that in other words, agents display the right or relevant kind of skill during cognitive practices because the environment in which their actions unfold actively takes part in the cognitive process, offering itself as *performatively qualified.*

Now, minimal representationalist theories, at a *general level*, seem to agree with this broad picture of practical cognition. Nevertheless, they are critical towards the attempt to explain *any* cognitive phenomena by endorsing a non-representational conception of practical knowledge. Indeed, if we focus on the

---

Epistemological Relevance Problem, which is the problem of how a system knows what is relevant in a defined cognitive task. Later, when I will refer to contextual relevance, I will refer to what has been called *Generalized Relevance Problem*, namely the problem of selecting (i.e. framing) only relevant information distributed in the context of a cognitive practice (Chow 2013: 312- 315).

relational aspect of this ability, a fundamental, problematic aspect can be shown. If it is true that practical knowledge is strictly dependent on the fact the world shows itself in a certain way, what about cases in which the world does not clearly show up as endowed with this or that "performance quality"? This is to say, what about cases in which the context of the cognitive practice is less regulated, unruly, or "noisy"? If the core claim of fully-non-representational accounts of practical knowledge is that agents display a regular and meaningful behavior in situated cognitive practices because the structure of the environment *directly* affords the activation of this or that skill (and this means that the agent's behavior is regulated by the environment), it seems to be the case that, if in those complex and unruly contexts agents still display intelligent behavior, their actions should be regulated by something else, by something that is not in the environment, but rather is in their heads: action-oriented representations.

To be accurate, it is worth to notice that, if we take into account the specific debate about the nature of practical knowledge (representational vs. non-representational), we will not see Andy Clark performing the role of the main character. Indeed, the discussion about whether practical knowledge, abilities and skills conceived as non-representational account for the issue of coping with the context or if we still need some kind of representational mediation interested other leading figures of philosophy of mind and cognitive sciences: Michael Wheeler (Wheeler 2005; 2008b; 2010a; Cappuccio, Wheeler 2012) and Hubert Dreyfus (Dreyfus 2002a; 2002b; 2008). This is a specific debate that has a Heideggerian flavor, and it has been developed around Dreyfus' attempt to think of a "Heideggerian AI", namely to a way of doing artificial intelligence aimed at producing models of on-line thinking that account for our practical, non-intellectual and non-intellectualized "being-in-the-world"[51]. Nevertheless, I think that there are some good reasons to consider this problem in relation to the broader literature about the enactive approach to cognition and the EMH.

First, these two approaches to cognition confronted in the debate, Wheeler's one and Dreyfus' one, are theoretically connected to the enactive version of Embodied Cognition and the Extended Mind. Considering

[51] Dreyfus refers to Rodney Brooks and colleagues' project developed the MIT Mobile Robot Laboratory. The idea of this project can be summed up by the slogan "Fast, cheap and out of control!". Indeed, the aim of Brooks and colleagues' work was building robotic creatures (Mobots) able to display intelligent and adaptive behavior without being controlled by a rich centralized representational system, considered to be not cheap for both theoretical and practical purposes (design). For a more detailed explanation of the project, see Brooks 1989; 1990; 1991; 1995; 1998.

Wheeler's approach, it is easy to individuate an affiliation with Clark's thought, because Wheeler himself explicitly endorses some core ideas of the EMH (e.g. Wheeler 2010b, about mind extension and extended functionalism). In contrast, Dreyfus' case is more complex. Indeed, Dreyfus does not make use of the label "enactivism". Nevertheless, his embodied approach to cognition centered on practice and action is not at odds with the core ideas of the enactive perspective, according to the interpretation of it I gave till now.

Second, the problems at the core of this debate are considered in the broader discussion about AORs by scholars who endorse an enactive approach to cognition and the mind (Rietveld 2008a; 2012a; Gallagher 2017). Moreover, Clark himself (Clark, Toribio 1994; Clark, Grush 1999; Clark 1997a), when he discusses the role of active representations in cognition, frames the problem by taking into account the relation between skillful coping and representational formats. Then, it seems that not only Wheeler's discussion about AORs is an attempt to respond to the philosophical problem I took into account here (the frame problem), but also Clark's view of AORs has a lot to do with this issue.

After this discussion of the philosophical reasons that encourage some scholars to consider AORs, I take into account some arguments of the debate in more detail. This survey has three main aims.

First, it aims at understanding if we really need representations to account for skillful cognitive activities, here conceived as action-perception loops.

Second, it is an attempt to wonder whether those new approaches to representation in cognition are actually able to account for *action in representation*, answering both to the worries of orthodox cognitive sciences, which claim that there cannot be cognitive science without representations, and to the embodied/enactive view, which suggests cognitive sciences and philosophy of mind to seriously go back to action and practice.

Third, this exam of the role of representations in cognitive process would like to be a chance to think about the problem of mind-extension again, wondering whether AORs support the development of this theoretical project, or if they rather prevent explanations to come to this radical conclusion about the nature of the mind.

II.3 Cognitive experience is just being there...plus Action Oriented Representations. Clark's approach to AORs.

> "Ninety percent of life [...] is just being there; and we have indeed charted lots of ways in which the facts of embodiment and environmental location bear substantial weight in explaining our adaptive success. [...] [But] we should not be too quick to reject the more traditional explanatory apparatuses of computation and representation. Minds may be essentially embodied and embedded and *still* depend crucially on brains which compute and represent." (Clark, 1997a: 143; text into brackets added)

Andy Clark, despite acknowledges that most part of our lives is "just being there" (that is to say that our lives primarily unfold through the embodied, practical and smooth engagement with the world), claims that our "being there" has to be explained by making use of some classical concepts of orthodox cognitive science: computation and representation.

His core idea can be put in this way: our cognitive practices are embodied and embedded and, precisely *in virtue of* these features, we should claim that cognition unfolds relying on representations that account for embodiment and embeddedness. Those kinds of representations are called "action-oriented representations".

According to Clark, Action-Oriented Representations (AORs) radically differ from classic "chunky", explicit, symbolic representations because, instead of

i) being representations whose key contents are tokenable strings of symbols, operated upon by a 'read/write/copy' architecture (Clark, Toribio 1994: 403),
and instead of
ii) re-presenting (i.e. mirroring, or accurately describing) properties of the world,

AORs *simultaneously describe* aspects of the world and *prescribe possible actions*. They are neither passive, crystallized pictures of properties of the external reality, nor pure control structures. On the contrary, they are poised between these two "cognitive slices" (Clark 1997a: 49). Moreover, they are described as "local", "action-dependent", and as reflecting "the profound role of bodily motion [...] in shaping and simplifying the information-processing problems to be solved" (Clark 1997a: 149).

To understand what Clark means by action oriented representations better, I think that it is worth to start my discussion trying to understand what prescriptive and descriptive functions of AORS are. To consider this point, it could be interesting to look at one of the sources that Clark considers in his discussion (Clark 1997a: 50; 238, footnote 5): Ruth Millikan's Pushmi-Pullyu Representations.

Millikan describes Pushmi-Pullyu Representations (PPRs) as representations that, despite are not the sum of a descriptive function and of a directive function, face both those ways at once. To make this point more clear, I give a closer look at the meaning of "descriptive" and "directive", by considering the example that opens Millikan's article *Pushmi-Pullyu Representations* (Millikan 1995: 185-186).

Consider a list of groceries. This one can be used for two distinct purposes. It might be used as a shopping list, which tells what to buy, or it can be used as an inventory list, which tells what has been bought. The list, considered in the first sense, has a directive function: it tells us what to do in order to accomplish a task. Moreover, it makes a normative claim (in the sense explained in §I.7) about the world, according to which the world is supposed to conform to the representation: if the shopping list does not match what the grocery bag contains, it is what is in the bag that is at fault. On the contrary, if the grocery list is used as inventory list, the function the two elements of the normative relation have is inverted: the representation is supposed to conform to the world, namely if the list does not match what the grocery bag contains, it is the list that is at fault. Indeed, the shopping list used as an inventory list describes what is the case in the world.

PPRs have both these two dimensions: they tell us what to do, by saying what is the case.

Another example of PPR is the following one (Millikan 1995: 190-192): the food call of a hen to its brood, whose function is to make the chicks coming to the place where food is. What is the structure of this very primitive representation? The call is evidently directive. It says something like "come here and eat". On the other hand, the representation has in itself the *condition* for the successful performance of the task (coming there and eating food). Indeed, when it directs action, it simultaneously says "here's food now". It gives information about the spatial location of the food, by describing a state of affairs. What is interesting is that such a representation does not merely mirror external reality, but by saying what is in the world, affords a specific kind of action. It connects directly with action because its nature is action-oriented. This goal-directedness makes the representation vary as a direct function of a certain variation in the environment, "directly translating the shape of the

environment into the shape of a certain kind of conforming action" (Millikan 1995: 190).

According to Millikan, pushmi-pullyu representations are also interesting conceptual tools to explain human cognition. For instance, intentions could be considered to be PPRs. Millikan considers intentions as internal states, as internal representations that have a directive nature (viz. they cause a certain kind of behavior). Nevertheless, she claims that explaining them only as directive is misleading. Indeed, it is commonly accepted that a person cannot actually intend something without believing that she will do that action. Then, if intention implies having a belief that P, then these kinds of representation have also a descriptive nature: the representational content of the belief describes what will be done.

Another example of PPRs are *perceptual representations*. Those ones are mental representations that map variations of the agent's perceived world by encoding those variations as possible perceptual actions: they map "variations in goals directly onto the represented future world" (Millikan 1995: 192). This is to say, perceptual pushmi pullyu representations function as proxies of future perceptual situations: they represent how the world will look like when the agent will act upon it.

This last example of pushmi-pullyu representation is particular interesting for the purposes of this paragraph. Indeed, it seems to me that what Clark calls "Action-Oriented Representation" is precisely what Millikan calls "pushmi-pullyu perceptual representation". To understand this point, it is worth to have a look at the first example of Action-Oriented Representation that Clark gives in his book *Being there*: Maja Mataric's work on mobots.

At the MIT Artificial Intelligence Laboratory, Mataric and colleagues worked on a project in AI aimed at designing a neurobiologically-flexible spatial representational model, which scientists implemented and tested on a physical autonomous mobile robot.[52] Mataric's mobot works in this way (Clark 1997a: 47 and ff.). It uses a set of quasi-independent layers, and each one constitutes a process route from inputs to outputs. Each layer works on a specific part of the environment. One generates boundary tracing (the walls the robot follows while it avoids obstacles); another one detects landmarks, registered as a combination of the robot's motion and its sensory inputs, and a third layer uses this information to produce a map of the environment. Those representations constitute what is called a cognitive map. This one is made of a network of landmarks, which are a combination of motor and sensory readings. The nodes

[52] See Mataric 1991 for a detailed description of the project.

of the map process information in parallel: an active node excites the nodes in the area nearby, generating expectations of the next landmarks that will be encountered in the map.

The basic idea of this model is that the cognitive maps the robot makes use of are made of nodes that combine *descriptive* information about the robot's own *movement* with *perceptual information*: this makes the map working as a controller for the robot's action. Moreover, by mapping the relation between information about the robot's movement and perceptual information about the environment using the propagation of signals among nodes, the cognitive map generates plans for real movement, making the robot able to react to real-time environmental conditions.

If I understand the description of Mataric's robot that Clark gives well, I think that it is possible to claim that the robot's cognitive maps are very similar to what Millikan describes by the expression "perceptual Pushmi-Pullyu Representations. Cognitive maps have this function: by generating plans for real movements, they *internally simulate*[53] what should be done by the robot in the environmental space, making it ready to cope with changes (e.g. angles, obstacles, and so on) in its perceptual real-time space. Then, the idea is that the robot successfully copes with its environment because it internally represents what it can do during the engagement with the real world, encoding perceptual signals as *possibilities* or "*orders*" for action.

Now, action-oriented representations in human cognition, according to Clark, are really close to Mataric's cognitive maps. They are *internal*

[53] It is worth to notice that Clark does not use the concept of simulation here. Nevertheless, it seems to me that "simulation" is a concept that fits with the general theoretical frame in which Clark discusses representations. For example, in the article he wrote with Grush in 1999 (Clark, Grush, "Towards a Cognitive Robotics", *Adaptive Behavior*, 7(1), pp. 5-16), which foreshadows many of the ideas developed by Clark in more recent publications, the example given in order to defend a form of minimal robust representationalism is the *emulator* one. An emulator is a mechanism (circuitry, software routine, and so on) that takes information about the starting (or current) state of a system and about the control commands issued as its inputs, and then gives a prediction of the next state of the system, by representing it as an output (Clark, Grush 1999: 4). This prediction sounds like a simulation of the state in which the system will be when it will do something according to the motor commands issued. This interpretation of the simulative power of those representations is not aimed at claiming that emulator theories and simulation theories conflate the one on the other. As Grush explains (Grush 2004), simulation theory and the emulator one differ because the first one usually claims that motor commands are just simulated, the latter claims that those controls are executed. Nevertheless, it seems to me that in Clark's explanation there is a simulative aspect: the representations is a proxy of states that action in the environment will provide to the system. This interpretation of the simulative aspect of those representations can be also supported by considering Clark's more recent work on Predictive Processing. For instance, in Clark 2013b, he talks about action planning and action-selection in terms of simulation ("simulations that allow us to explore possible course of future action"), and he mentions Clark, Grush 1999 in this discussion (Clark 2013b: 1-2).

personalized representational states, conceived as *neural encodings*, which map idiosyncratic, locally effective features to guide behavior (Clark 1997a: 151).

Clark considers them to be the most evolutionary and developmentally basic kinds of representations (Clark 1997a: 152) because they seem to be at the core of humans' fundamental and primordial cognitive activity: reacting *selectively* to environmental stimuli, which are *complex* and *unruly* (Clark, Toribio 1994: 419). This allows the agent to display the right kind of behavior in a given situation.

AORs are defined as:

a) *action-specific* because they are tailored to the production of the specific behavior required (they are the mental antecedent or a simulation of an action performed on-line, in the real world);

b) *egocentric*, because they encode features of the environment in a way that accounts for the robot's history of sensorimotor experiences, namely features of the environment are represented as intertwined with memories of bodily motions;

c) *intrinsically context-dependent*, because context is "woven into the representation-using mechanism's basic operating principles" (Wheeler 2010a: 326). This is to say that AORs co-vary with external states, explaining how something inside the agent is about something outside the agent (Chemero 2009: 50).

Then, AORs seem to be perceptual pushmi-pullyu representations "neurally located". They encode perceptions as motor commands tailored to selected features of the environment: they are said to guide action because they represent parts of the environment as cognitive maps endowed with a *conative power*[54]. Indeed, like Millikan's representations, those internal descriptions of selected parts of the perceptual array have a "you must do" nature, and they are endowed with this feature because of their s*emiotic structure*. They are *indexical* or *deictic* representations[55]. They are entities that relate specifically to the agent and they have a functional value because they play a specific role in the activity the agent is engaged in. They are not objective, they are not tokens of a symbolic type, but they rather point to different situated objects (adapted from Agre 1997: 243). By referring to those specific objects as they are perceived by the agent, they make her ready to react, to perceptually engage

[54] This is a liberal use of the expression "conative aspect of pragmatic representations", found in Nanay 2013: 20.

[55] This is my development of Clark's short explanation of AORs, motivated by the connection between AORs, Millikan's pushmi-pullyu representations and Agre's deictic representations that he individuates (Clark 1997: 152).

with those objects without making complex operations of subsumption of the singular represented item under the concept that explains it. AORs are perceptual in the sense that they represent this or that individual object, with its perceptual features, and specific sensorimotor commands, which are constrained by the situation in which the object represented is embedded, and by the history of the agent's previous situated experiences.

That is why those internal representations are said to be *embodied* and *embedded*[56]: "they stand for what's happening to me, right here, right now" (Chemero 1998). They are embodied because they represent perceptual stimuli as tied to the agent's sensorimotor experience, and they are embedded because the prescriptive representation of motor commands also encodes information about the local, specific environment, where by the expression "local environment" I refer to the agent's peripersonal space. Moreover, they are *active*: the content of those representations activates a disposition towards embodied actions.

At a general level, it seems possible to claim that the reason why Clark talks about mental representations as embedded, embodied and action-specific is the one taken into account at the beginning of my discussion about second-wave cognitive sciences, in §I.4: the symbol grounding problem. The argument pro "grounded representations" was the following one: in order to account for cognition as a situated and active process, scientists should i) put action, ii) the body, and iii) the specificity of the context in representational mechanisms. In doing so, according to Clark's approach to cognition, representations should be also consistent with an extended explanation of the mind, precisely because mental representational activities are conceived as dependent on parts of the cognitive machinery that are not embedded within the boundaries of the skull.

More specifically, the way Clark thinks of action-oriented representations seems to be an attempt to get rid of any highly intellectualistic and abstract conception of cognition, project that characterizes his philosophical production.[57]

Nevertheless, despite this attempt to get rid of the intellectualism of classical cognitive sciences, Clark's endorsement of AORs seems to hide a bit of *conservationism*, whose reasons, to me, are not always clear to understand.

Indeed, looking at his book *Being there*, it can be found out that AORs have been introduced in Clark's explanation as a critique to Gibson's theory of direct perception. About this point Clark says:

[56] See Rupert 2009: 200 for the expression "embedded representations".
[57] See for example Clark 1997 Chapter 2; Clark 2016a Chapter 4; Ward, Roberts, Clark 2011.

> "A related view of internal representation was pioneered by the psychologist James Gibson (1950,1968,1979). This work made the mistake, however, of seeming to attack the notion of complex mediating inner states *tout court.* Despite this rhetorical slip, Gibsonian approaches are most engagingly seen only as opposing the encoding or mirroring view of internal representation. Gibson's claim, thus sanitized, was that perception is not generally mediated by action-neutral, detailed inner-world models. It is not mediated by inner states which themselves require further inspection or computational effort (by some other inner agency) in order to yield appropriate actions. This is not, then, to deny the existence and the importance of mediating inner states altogether" (Clark 1997a: 50).

If I understand well what Clark says, it seems that AORs should be considered to be pervasive epistemic posits, in the sense that *any* perceptual episode should be said to be guided by those internal structures.

Nonetheless it is not actually clear to me *why*, according to Clark, we need action-oriented representations to account for the relation between action and perception. Indeed, Clark only says that the attack to the notion of internal representation *tout court* is just a rhetorical slip, and later he says that the anti-representationalism of what he calls "Radical Embodied Cognition" (REC) is unwarranted and counterproductive, because "it invites competition where progress demands cooperation" (Clark 1997a: 149). Later, to provide some arguments to justify the claim that anti-representationalism is unwarranted, he states that REC rejects representations *tout court* because:

i) it sticks to a narrow concept of mental representation (i.e. amodal, chunky, symbolic and explicit), without considering the possibility of explaining cognition as a continuum of representational degrees (see also Clark, Toribio 1994);

ii) it is concerned with cognitive phenomena that are not *representation-hungry* enough (Clark 1997a: 149) - namely cases that involve simply physically present and simply specifiable parameters" (Clark, Toribio 1994: 412)- and inferring from this that cognition, in general, does not unfold through internal representations.

Considering the first point, it can be said that this claim could sound true only if we consider *Being There* only. Clark has been one of the first scholars to introduce the term "action-oriented representation". At that time, the debate about AORs was not widely developed, so anti-representationalist approaches to cognition Clark refers to in 1997 actually reacted to classical views of

representations, and not to embodied-embedded representations. Obviously, while the debate went on, objections directed specifically against the concept of AORs have been raised. I will consider these objections in the following paragraphs.

The second point, namely the idea that REC considers only cases that are not representation hungry enough and *then* endorses an anti-representationalist approach to cognition is trickier, and it needs to be taken into account more seriously.

As previously said, in *Being there*, Clark develops the concept of action-oriented representation to reassess in a representational fashion Gibson's idea that perception is an activity that takes place through the relation between an embodied and moving agent and her environment. Hence, as I said before, AORs seems to be an occasion to state that perceptual experiences, that is cognitive processes usually considered as low-level cognition, *always* unfold relying on AORs. The on-line, real-time, and active cognitive phenomenon of perception requires the mediation of internal representations in order to be explained. Indeed, according to Clark, the idea that perception is direct (i.e. non-representational, non-inferential, not internally regulated) is misleading: claiming that perception is direct is just a rhetorical device aimed at dismissing the idea that perception lays in the realm of passivity.

To me, the idea of *representation-hunger* makes Clark's endorsement of Minimal Robust Representationalism (Clark, Grush 1999) - namely the idea that our "being there" should be explained by relying on action-oriented, egocentric, and context-dependent representations – *unclear*.

Indeed, the idea of representation-hunger seems to implicitly suggest that *some* cognitive situations (although they are probably many, according to Clark) require internal representations to be explained; other kinds of situations do not.

According to Clark (Clark, Toribio 1994: 419) representation-hungry problems are cases that involve one or both these two conditions:

i) reasoning about absent, non-existent, or counterfactual states of affairs (Clark, Toribio 1994: 419) – and these cases seem to include phenomena such as imaging and remembering;

ii) the cognitive task requires the agent to display a "selective sensitiveness", namely to select parts of an environment whose physical manifestations are complex and unruly (Clark, Toribio 1994: 419) or, more roughly, it is a case of "undetermined perception" (Mandik 2005: 291).

In my discussion, I do not take into account the first kind of cases. This because the debate between representationalists and anti-representationalists concerning imaging or reasoning about the absent, at present, is not very developed (although some anti-representational explanations of those phenomena has been offered in Hutto 2015; Degenaar, Myin 2014; Gallagher 2017; Rooij, Bongers, Haselager 2002), and also and especially because the problem I am dealing with now (representational vs. non-representational accounts of active perception) does not concern point i) but point ii).

An example that can be helpful to consider point ii) is that about the role of internal cognitive resources that play the function of effective *tracking devices*[58], namely representations that occur when an agent is continuously coupled with the object of her tracking (Clowes, Mendonça 2016: 31) when the physical manifestations of the information the cognitive system needs are widely various (Clark 1997a: 167) and difficult to detect.

For instance, think about the cognitive task of picking up all the valuable items in a room (Clark 1997a: 167).

Clark does not discuss this example at length, but I will try to develop it further because the case of responding selectively to environmental stimuli is a crucial point in the debate about AORs: it is precisely on this point that affordance-based theories of perception work on, discussing the representationalist position in a critical fashion. The lack of clarity of the arguments offered by representation-friendly explanations, I claim, is an obstacle to understand the overall project of action-oriented representations because it undermines the construction of a ground of confrontation among theories. To put in other words, it seems to me that approaches that endorse AORs make objections to eliminativist views, claiming that it is not clear how agents react to some perceptual stimuli and not to others in a given context, but when they try to offer an alternative view, they are very vague. This makes difficult to understand if AORs are good epistemic posits in the case taken into account or if they do not play an interesting epistemic role.

So, suppose that an agent has to pick up all the objects designed for drinking in the kitchen. The environment where the cognitive task is complex because the room is full of objects designed for different purposes. Moreover, despite the agent should focus exclusively on those objects designed for drinking, she receives a lot of sensory inputs; therefore she has to discriminate between inputs that are valuable for the ultimate aim of the task she is engaged in and

[58] This is close to what Ramsey critically calls "receptor notion", namely the idea that neural structures are said to be representations because they do the job of detecting parts of the environement, to the extent that they respond to given environmental or bodily conditions in a reliable manner (Ramsey 2007: 120; Clowes, Mendonça 2016: 31).

those that are not primarily interesting. Moreover, the objects designed for the action of drinking show up according to different physical manifestations (they have different colors, their shape vary a little bit, and so on). In this case, Clark would say that, because of the complexity of the environment (the agent's perceptual target of attention is elusive because it is embedded in a context full of noisy information), the agent has to internally represent selected parts of the environment (all the objects for drinking).

This would

i) provide the agent with an *internal target of attention* that will guide her behavior, by making her focused on some perceptual stimuli only;

ii) make her ready to act in a certain way (e.g. it would make her ready to grab objects for drinking and not ready to throw them) because the AORs represent the objects for drinking under the respect of "grabbing" or "picking up" and not under the respect of "throwing". The internal action-oriented representation would say something like "there/object/for/drinking/-grab/it!"[59].

Now, it seems to me that the case considered above is a good example to think about Clark's argument pro AORs. In this example, the environment where the task takes place is considered to be informationally unfriendly or noisy because it offers the agent a wide variety of perceptual information. This is said to make an explanation of adequate action difficult if the explanation does not suppose internal proxies of the external objects the action should be performed upon.

To me, it is particularly striking that the situation where the cognitive task takes place is defined as unfriendly, and then as "representation-hungry". Indeed, the cognitive task Clark takes into account seems to be just an example of what agents do in their everyday cognitive experiences, which do not take place in defined "experimental settings", but in complex and informationally varying environments. Thus, if this point is taken into account, and if this cognitive task is still defined as "representation-hungry", it would be more

[59] Described in this way, Clark's AORs (and also Millikan's PPRs) sound like an attempt to find a solution both to the Epistemological Relevance Problem and to the General Relevance Problem. Indeed, the prescriptive function of AORs, according to which representations prescribe a specific kind of action, selects relevant information in the environment (thanks to the indexical or deictic aspect of the representation) and the right kind of skill to display (the representation says what should be done) all at once.

accurate to say that human action-perception loops always require the mediation of action-oriented representations, and this would be consistent with Clark's purpose of reassessing Gibson's theory of perception in a representational fashion.

Nevertheless, if I understand Clark's argument well, this is not the case. In fact, cases like this are taken to be *specific* representation-hungry problems and the idea of representation-hunger is introduced to demonstrate that eliminativist explanations are wrong because they consider problems that are not representation-hungry enough (Clark, Toribio 1994: 418). The problem seems to be this one: some of the cognitive episodes Clark takes into account to demonstrate that they are representation-hungry seem to be the same cases he thinks to be not representation-hungry enough when he makes his objections against eliminativist approaches to representation.

Therefore I think that two possible philosophical scenarios can be envisioned.

i) We could claim that any cognitive experience requires representations in order to be explained, and then we should dismiss the concept of representation-hunger. This entails that we should find other objections to theories of direct perception, which do not entail the concept of representation-hunger.

ii) Or we could not use the notion of representation to account for perceptual experiences, preserving -at least at a provisional level- the idea of representation-hunger for the first kind of cases (reasoning about the absent or counterfactual) only.

As it will be explained in my exam of other objections against AORs, I am prone to think that option ii) is more viable. In the next paragraph, I will take into account those objections to AORs. Then, I will make my considerations about AORs in the Extended Mind Hypothesis framework more explicit.

II.4 Representational criteria. Peircean interventions.

A useful source to consider the problem of action-oriented representations more in depth is Gallagher's article *Are minimal representations still representations?* (Gallagher 2008a)[60]. To be accurate, it is worth to notice that Gallagher's article focuses more on other versions of action-oriented representation that show up in the "4Es" debate, in particular on Wheeler's and Rowland's ones, and on Clark and Grush's emulators[61]. Nevertheless, I think that some of the objections to minimal forms of representationalism that Gallagher makes can be raised against Clark's approach to AORs in *Being there* too. Indeed, this approach preserves some of the aspects of the emulator notion, critically considered by Gallagher (see §II.3 of this work). Moreover, many aspects of the notion of AOR developed in *Being there* are consistent with the notion of action-oriented representation offered by Wheeler in *Reconstructing the cognitive world* (Wheeler 2005). Both versions of action-oriented representation are said to encode the external world in terms of possibilities of action (Wheeler 2005: 197-199; Clark 1997a: 149), to be poised between mirroring and control functions (Wheeler 2005: 197; Clark 1997a: 49), to be the result of natural selection for adaptive purposes (Wheeler 2005: 198; Clark 1997a: 152), to be heavily context and body dependent (Wheeler 2005: 197; Clark 1997a: 149), and to be strictly tied to the individual agent (Wheeler calls them "egocentric", Wheeler 2005: 197; Clark calls them "personalized", see Clark 1997a: 149). The most significant point the two philosophers disagree on is the spatial location of AORs. In fact, Wheeler says that AORs are distributed across brain, body and environment (Wheeler 2005: 221-222; Gallagher 2008a: 357); on the contrary Clark, in *Being there*, considers those representations to be internal, since he identifies those structures with neural encodings. For example, he says that "to the extent that the biological brain does trade in anything usefully described as 'internal representation', a large body of those representations will be *local* and *action-oriented* rather than objective and action-independent" (Clark 1997a:149).

The philosophical strategy that Gallagher follows to understand whether representation is necessary *in* action, namely to understand if representation is a crucial and unavoidable part of action itself (and this means investigating perceptual processes too, since in the debate I am taking into account

[60] Notice that this article is just a theoretical reference point for the structure of my argumentation, which, later, I develop autonomously.
[61] For a broader view of those approaches to AORs see Rowlands 2006a; 2006b; 2012; 2015; Wheeler 2005; 2008a; 2008b; 2010a; Clark, Grush 1999.

perception is considered to be an activity[62]), is particularly interesting. Indeed, instead of considering individual cases philosophers refer to, he focuses on the concept of representation in itself, in order to understand if AORs meet the criteria required to talk about representations. This is aimed at questioning the theoretical status of AORs in themselves; that is why Gallagher's argumentation can lead to strong conclusions about AORs. If the comparison between the widely accepted criteria to talk about representations in cognitive sciences and the concept of AOR will demonstrate that AORs are not actually representations, this will show that i) talking about AORs is just an arbitrary terminological choice, and that ii) explaining the cognitive phenomena taken into account by appealing to AORs is philosophically unwarranted.

I begin my discussion about AORs by making "representational criteria" explicit.

Although in philosophy of mind and cognitive sciences there are many notions of representation, and there is no consensus on their nature (Ramsey 2007: xi), the majority of scholars seems to agree on some elements that define the *classical* notion of representation. Those criteria are:

i) Internality: representation-tokens exist in clearly identifiable regions of space. In philosophy of mind, the boundaries of those regions of space are those of mental images, syntactically structured symbols, neural configurations, or any internal configuration of the subject. In this sense representations are spatially discrete.

ii) Genuine duration: representation-tokens have identifiable temporal boundaries. Occurrent mental states have representational contents. They are activated, namely their status changes from dispositional to occurrent, when a representation is tokened, viz. when representational contents are brought on-line by some capacity. In this sense, representations are temporally discrete.

iii) Standing-for: Representations are proxies (they stand-for) of something else (e.g. perceptual stimuli, objects that fall under a category), and they do so because they bear a content.

iv) Interpretation: representations mean nothing in themselves. In order to acquire meaning, they have to be interpreted by someone (e.g. an interpreter, in the case of words and sentences) or something

[62] As previously sketched out, this is particularly clear and widely developed in sensorimotor versions of enactivism (Noë 2004; Noë 2010; O'Regan, Noë 2001); nevertheless, also "extended" approaches to perception are aimed at accounting for perception as active (Clark 2016a, in particular Chapter 4; Clark 2008, in particular §1.3).

(e.g. an interpretive function that specifies the place that a representation occupies in the subject's representational economy).

v) Passivity: classical representations are passive. To the extent that they acquire their meaning in virtue of the job of an interpretive function, their capacity to stand-for X depends on what is done with those representations by the interpretive function. Representations do nothing, but something else does something with them. They are the media of an activity, not an activity in itself.

vi) Decouplabiliy: an item is genuinely representational if it is decoupleable from its environment, namely from the states of affairs that it represents. This means that, if R is a representation and T its target, a) R and T are not in constant causal contact; b) R represents T when it is absent; c) in doing so, R has an adaptive function.[63] This condition is tied to the epistemic function of representations in the explanations of human behavior. In fact, representations have been introduced in cognitive sciences to explain the mechanisms that guide agents' behaviors in the absence of the feature they represent (Haugeland 1991: 172). If an item is not decoupleable, then it is not clear to what extent it plays the role of a mediation function. An item that plays a certain function only when it is in contact with the immediate environment (Rowlands 2006a: 46) seems to entertain just a causal relation with the environment, or with parts of it.

These six conditions to talk about mental representations in a classic and orthodox way are not endorsed as a whole by "second-wave" cognitive scientists.

First, considering the broad picture of "4Es", it can be shown that the "internality condition" is sometimes taken more *loosely*. For instance, consider what has been said about external representations in §I.8. The example I gave was that of the Tetris zoids, which stand for the bricks of the wall the players deal with. This kind of representation -like other kinds of external representations, such as tags, material symbols, written words, tokens of mathematical notations - is not located within the boundaries of human heads. In this sense, to mark the contrast with the classical notion of representation, such a representation is said to be external. Nevertheless, those representations seem to still meet, to some extent, the criterion of internality, if the word "internality" is meant to refer to the fact that representational tokens are

[63] This passage is taken from Chemero 2009: 48.

embedded in a representational system[64] in which, thanks to their relation with other representational items, their spatial location (namely their being spatially discrete items) is shown.

Second, the "passivity condition" is sometimes considered to be unimportant. As already explained, the idea that representations do not just undergo action, but rather "*do* various things" (Ramsey 2007: 18) is a consequence of two more general theoretical points. i) The project of getting rid of intellectualism in the explanation of cognitive processes; ii) warranting embodied knowledge a special role in explaining cognition, and this means accounting for sensorimotor knowledge as the core of cognitive processes.

Third, the "decouplability condition" is critically questioned. To be accurate, it is worth to notice that not all the explanations that make use of AORs dismiss the idea of decouplability. For example, Wheeler explicitly rejects it (Wheeler 2005: 219), but Clark is not always clear about this point. In the article written with Grush (Clark, Grush 1999), the two authors want to preserve the decouplability condition. When they describe the job of emulators, they talk about the process of anticipation of states or stimuli, and they say that those anticipations are decoupled representations, namely representations of a future X, that is not there when it is represented. On the contrary, in *Being there*, when Clark talks about Haugeland's representational criteria (Haugeland 1991), he says that the role of decouplability is "overplayed" (Clark 1997a: 144). Clark's idea is that, if we consider the *strict* application of the decouplability criterion as an unavoidable requirement to talk about representations properly, we rule out the description of inner states as genuinely representational. This "seems unappealing in view of the very real explanatory leverage that the representational gloss provides, and it is also out of step with standard neuroscientific usage" (Clark 1997a: 145).

For example, he considers neural populations which are said to encode positions of the head. Those encodings are not decoupleable from what they represent: they are activated when they are in causal contact with what they represent. In this case, Clark claims, explaining this phenomenon by making use of a representational gloss helps us to understand how information flows within the system. Then, even if those codings do not fulfill the decouplability condition, we should describe them by making use of a representational vocabulary: "the lack of decouplability does not in itself seem to deprive the representational gloss of all explanatory force" (Clark 1997a: 238, footnote 4)[65].

---

[64] Notice that, according to this interpretation, the internality condition conflates with what is commonly known as "combinatorial constraint", namely the idea that an item counts as representational when it is part of a general representational scheme (Haugeland, 1991: 62; Rowlands 2009: 118).

[65] On the issue of decouplability in Clark, see also Clark 1997b: 472 and ff.

To understand whether the notion of action-oriented representation I am taking into account here is a genuine representation or not, it is worth to consider the set of conditions mentioned above more closely, in order to understand whether those criteria are *all* necessary to properly talk about representation. My hypothesis is that some of them are necessary, other conditions are not. I suspect that they have been introduced for "consistency" reasons, namely to fit with a specific approach to cognition, that is that one grounded on the "mind-computer" metaphor. To put it in other words, I claim that there are some core conditions that define representation in general, as a theoretical object; other conditions are required by some explanations of cognitive processes only.

To show what the conditions I consider to be fundamental to talk about representations properly are, I consider the etymology of the word "representation", in order to understand what it means, and to make clear the commonsense framework representations in cognitive sciences come from[66].

"Representation" comes from the Latin verb *re- + praesentō*. The word consists of the intensive prefix *re-*, which means "again", "once more", "anew", and the verb *praesentare,* which means "to show", "to exhibit", "to display", "to bring something to someone's mind", "to stand in place of"[67]. Therefore "to represent" means showing something again, standing in its place. Moreover, the verb "to represent" means "to act or speak on behalf of somebody", "to assume or occupy the role or functions of someone "[68]. Representations act or function as if they were the object or subject they stand-for. They do what the things they represent do.

By focusing on this commonsense understanding of representations, some core conditions to talk about representations (and notice that some of them are the same endorsed by the orthodox representational framework) can be individuated.

---

[66] At first glance, recalling the commonsense understanding of representations can look like an unwarranted philosophical move. Indeed, one can say that the notion of representation cognitive scientists deal with is a specific epistemic posit, our common use of the word "representation" cannot account for. Nevertheless, I think that this first impression is misleading, because, as Ramsey says (Ramsey 2007: 8), cognitive scientists seem to assume that they are tapping into a pre-theoretical understanding of representations.

[67] See The Concise Oxford Dictionary of English Etymology (http://www.oxfordreference.com/view/10.1093/acref/9780192830982.001.0001/acref-9780192830982-e-12741?rskey=fpzCsw&result=12742), and the Online Etymology Dictionary (http://www.etymonline.com/index.php?term=represent&allowed_in_frame=0).

[68] See The Oxford English Dictionary (http://www.oed.com/view/Entry/162991?rskey=6i4QWK&result=2&isAdvanced=false#eid).

i) Standing-for. Representations are items that, despite sometimes are isomorphic with what they represent (i.e. they have the same structure of the object represented), are different from what they represent. They are distinguishable items: representations are proxies of something else.

ii) Even if in the etymological definition of the word "representation" it is not explicitly mentioned, it seems that the "standing-for condition" entails the "interpretability" condition. Indeed, "to represent" means "acting or speaking on behalf of somebody". Words and actions acquire a meaning because they are words or actions for something or somebody, who interprets (even if this does not necessarily mean interpreting by means of an intellectual act) what is represented.

iii) Moreover, the standing-for condition seems to entail what has been previously defined as "genuine duration". In order to stand-for something, a representation has to be interpreted, and interpretative acts are something that takes place in time. The identity of a representation, namely what defines it as this representation and not that representation, depends on the meaning that this representation acquires, and meaning is acquired by means of an activity that happens in time.

iv) Even if the "decouplability condition" is not always considered to be necessary to talk about representations[69], I think that it is. Consider what said in point i) of my list. Representations stand-for something else, and this means that they are distinguishable from what they represent. The decouplability condition is strictly linked to the standing-for condition, central in our commonsense understanding of representations, and in scientific accounts as well. If an item or a process is not decoupleable from what it stands-for, it is not clear how identity boundaries between the two items or processes can be drawn (Rowlands 2009: 126).

What the commonsense meaning of the word "representation" does not entail is the fact that cognitive representations are internal to human heads and that they are passive.

The first aspect has been already taken into account when I discussed embedded and extended approaches to cognition. What was shown is the

[69] See Chemero, Eck 1999, and Chemero 2000.

fundamental role that those representations play in distributing the cognitive load of practice. This demonstrated how those external items shape and determine cognitive practices, by playing the same functional role that internal representations play in cognition. This demonstrated how the external location of representation does not prevent an item to be considered to be a component of cognitive processes.

The second aspect has to be examined more closely. The classical definition of representation described them as passive by appealing to the "interpretability condition". Representations were said to be passive because, taken in themselves, they mean nothing: they acquire their meanings thanks to interpretative acts, namely thanks to the activity of something that lies outside of the representation. As previously said, I agree with definitions of representation that endorse the "interpretability condition". Nonetheless, I disagree on inferring the passivity condition from the interpretability criterion.

Indeed, I claim that interpretability makes representations passive only if a specific model of representation is endorsed: the dyadic conception of representations.

Looking at the debate I am examining, it seems to be the case that cognitive scientists endorse a dyadic concept of representation. Indeed, they focus on two aspects: content and format. Representations bear a content (properties or relations of the environment) that is said to entertain a causal relation with behavior (behavior is caused by contents mentally represented), and they vehicle it in a certain way (e.g. by representing it in a modal or a-modal way). Content and format define the theoretical object "representation". The interpretative function is not said to be part of this representational entity in itself: that is why representations are said to be passive. What actually does something with contents (the interpreter or interpretive function) lies outside of the representational entity.

My point is that there is a viable way to preserve the interpretability condition, and to give it a central role in defining representations, without entailing passivity: conceiving of representations as *triadic relations,* as C. S. Peirce did in his work in Semiotics.

So far, this option has not been very popular, and this depends on the fact that C. S. Peirce's thought, and the classical Pragmatist tradition more in general, did not receive much attention in philosophy of mind and cognitive sciences. This underestimation of the pragmatist tradition in philosophy of mind is due to the fact that the core ideas about cognition endorsed by pragmatism seemed to preclude the access to philosophy of mind in itself, especially if "philosophy of mind" is a label meant to refer to critiques of the behaviorist explanatory paradigm (Daddesio 1995). As a matter of fact, the philosophical position endorsed by Pragmatism, commonly known as

"knowledge externalism" (Meyers 1999) or as "cognitive externalism" (Burke 2014; Aydin 2015), holds that cognition does not unfold through internal representations (namely ideas or concepts we have access to *via* introspective acts – W2: 193-210, from "Questions Concerning Certain Faculties Claimed for Man", 1868) but that the whole function of *thought* is to produce *habits of action.* This is to say that, to explain cognition and mental processes, pragmatists look at how and when habit causes us to act. "Mental" is a) the stimulus to action derived from perception -that determines the "when" of cognition-, b) and the purpose of action that causes a sensible result -and this is the how of cognition (W3: 265; from "How to make our ideas clear", 1878).

These pivotal ideas of pragmatism lead to describe the mind as something that is externally produced or realized by means of practice. This point is theoretically very far from what mainstream philosophy of mind and cognitive sciences claim. Indeed those ones conceive of the mind as constituted by mental contents, and define action just as a non-cognitive effect of mental processes. Pragmatism, by defining the mind as an effect of (conceivable) actions controlled by habitual practical knowledge, and then developing its explanation "out of the black box", completely knocks over the theoretical pillars of mainstream philosophy of mind: it is "Cartesianism read from right to left" (Fodor 2008: 12). According to Cartesian philosophy of mind and cognitive science, the pragmatist stance is "the worst idea that philosophy ever had", a "bad cold" philosophy should get rid of (Fodor 2008: 9, 11). According to "Cartesian philosophy", the priority of acting on thinking that Pragmatism endorses as its fundamental idea (which means holding that the distinctive function of the mind is guiding action) seems to be so naïf that it does not understand that our ability to act is grounded on a more fundamental level, namely on propositional knowledge, which makes action-planning possible. To plan an action, the subject should be able to represent how the world would be if that action would succeed, according to truth conditions (Fodor 2008: 13).

Given that the aim of my discussion is not to determine whether Fodor is right, neither my task is defending Pragmatism from those objections, I will not explore this point further. Nevertheless, my suggestion is to keep this strong criticism against pragmatism in mind, and to consider it as one of the reasons why it has not influenced cognitive sciences for many years[70]. I claim that this

[70] Notice that, in the last years, this trend changed. Indeed, the birth of action-oriented paradigms in cognitive science gave birth to interesting research focused on the theoretical value of the pragmatist tradition for cognitive sciences. Among others, see Aydin 2015; Burke 2008; 2014; Engel, Maye, Kurthen, Köning 2013; Fusaroli, Granelli, Paolucci 2011; Gallagher 2009; 2014; 2017; Jung, Madzia 2016; Menary 2007; 2009; 2016; Johnson 2006; Johnson, Rohrer 2007; Johnson 2010; Paolucci 2011; Roy 2013; Shook, Solymosi 2014a; 2014b; Skagested 2004; Steiner 2008; 2013; Tiercelin 1995.

precluded the chance to make some ideas clear in cognitive sciences and philosophy of mind: this the case of interpretability and the passive status of representations.

To take into account this point, I consider one of the core ideas of Peirce's theory of knowledge. According to Peirce, cognition unfolds through signs (see the article "Some consequences of four incapacities", 1868; W: 211-242). A sign is a representation (indeed Peirce writes that "'representation' and 'sign' are synonyms"; CP 8.191), and the word "representation" refers to "an object which stands for another so that an experience of the former affords us a knowledge of the latter" (Peirce, MS 389, 1873). According to Peirce, representations involve three *relata*: a) a Representamen, which is the first correlate of a triadic relation (MS 1345) and that in the contemporary terminology would be called "sign-vehicle", b) its Object, the second correlate, c) and a third, which is called "Interpretant" (CP 2.242), or sometimes "interpreting act" (CP 2.230).This last one, by entertaining with the object the same relation that the sign-vehicle entertains with it, brings forth new knowledge about that object. To put it in other words, signs bring forth a relation between a First, a sign-vehicle, and a second, the Object, by means of a Third "that looks inside of the First through a second" (translated from Fabbrichesi 1986: 97).

For example, think about symbolic signs, namely general signs (CP 1.558) that refer to their objects in virtue of a law (for example an association of general ideas; CP 2.249). Take a term (CP 2.25) like "cat", for instance. The written or uttered word "cat" is the sign-vehicle (first correlate), the cat is the object (second correlate), and the meaning of the word "cat", the idea of cat, is the Interpretant (third correlate). According to Peirce, the word "representation" refers to the dynamics that connects these three *relata*. Representation is: the relation between the object and the sign-vehicle, which is determined by the object (MS 793), and the relation between this relation and the Interpretant, which is actively determined by the sign vehicle, in the sense that the sign-vehicle affords an Interpretant (MS 793). The sign-vehicle represents the object it stands-for under a certain respect and the Interpretant brings into light this relation by representing it. In doing so, the Interpretant shows the meaning the sign-vehicle is meant to carry.

The notion of Interpretant is crucial in this explanation. The Interpretant is said to fulfil "the office of an interpreter, who says that a foreigner says the same thing which he himself says" (CP 1.553). It should be noticed that, considering the broad picture of Peirce's semiotics, this linguistic description should not be interpreted literarily (Deacon 2014: 97). Indeed, not only Peirce distinguishes the Interpretant from the human Interpreter, who translates one sign into another sign (see the letter to Lady Welby of March 14th, 1909; SS: 109-10),

but he also distinguishes different kinds of Interpretants. The *meaning* of a triadic representation (MS 318:163; 37-38), can be a mental effect, an emotion, a feeling, an effort (MS 318: 40-1; CP 5.475), an imperative command (CP 5.473), an affordance for action.

This variety of Interpretants depends on the *situated* nature of semiosis (Atã, Queiroz 2014; Fusaroli, Paolucci 2011: 18-19; Queiroz, Merrell 2008; Violi 2008: 252), namely on the temporality (that is on the stage the process of semiosis is at; Deacon 2014: 97) and on the materiality of the semiotic relation.

To explain this point, it is worth to take into account an example offered by Peirce between 1905 and 1907, in a letter-article addressed to the editor of the journal *The Nation*. Consider an agent who is listening to a piece of music. The piece of music represents iconically (Santaella 2015; Short 2007: 204) –namely in virtue of its own internal qualities (CP 2.92)[71], like perceptual qualities, for example- the composer's musical ideas. Those ones, since they are perceptually represented, at first are interpreted by means of the sentiment or feeling the piece of music excites in the listener (CP 8.335) when she perceives successions of musical notes. This feeling, although it cannot constitute a complete and ultimate comprehension of the composer's musical ideas, is an Interpretant, an *emotional Interpretant* of the sign (CP 5.475). The listener's feelings/emotions at the individual time T, in the context C, are a "first" piece of knowledge about the composer's musical ideas, they are an initial step in the process of knowledge. Those feelings, on their hand, cause some further reactions or effects (Short 1996: 509). For example, they can entail a bodily effect (CP 5.475, where Peirce talks about "muscular effects"), such as changes in the bodily posture of the agent. According to Peirce, this is an *energetic Interpretant* (it is called "energetic" because it entails a mental or bodily effort), whose cognitive effect is focusing the agent's attention (MS 318: 443-45) on that individual occurrence of feeling triggered by the piece of music. The Interpretant of the sign changed, namely new cognition has emerged by means of an energetic Interpretant because the *situation* of semiosis changed. The emotional Interpretant afforded new knowledge about the object, becoming the first correlate of a new representational relation, and affecting the energetic Interpretant by means of the action of a new sign-vehicle. The new Interpretant, the energetic one, is an individual act, hence it cannot work as a sign for the generalized experience of music. That is why Peirce postulates a third kind of Interpretant, more developed than the others, and able to push forward cognition towards a more general level: the *ultimate logical Interpretant*. The ultimate logical Interpretant is the effect of a series of signs, that is the repetition of the experience of music. Thanks to generalizations of acts of response, it

[71] On this specific point about iconicity see also Fabbrichesi 2017: 57 and ff.

changes the agent's general way of acting (Peirce describes this as "*habit-change*"; CP 5.476), namely the habit-based actions performed when that piece of music is played. The development of signs produces a modification of the experience of music (Eco 1976: 1465-1466). The agent will be ready to act in a certain way (for example she will consciously appreciate her feelings for music, she will enjoy the structure of the piece of music, focusing her attention on it and not on the individual notes, and so on) when she will listen to that piece of music again.

What is particularly interesting in this explanation of representations is the way Peirce conceives of them as *active*. The reason that accounts for this active aspect of representation is twofold.

On the one hand, by conceiving of the Interpretant -namely the effect of meaning of a representation- as a constitutive part of the representational relation itself, Perice's explanation avoids passivity. Representations stop to be entities that should be manipulated in order to achieve meaning, but they are defined as the process by means of which sign-vehicle, object and meaning are connected. Meaning arises in this relation, and it is *internal* to the representation itself: the activity of the Interpretant is the third correlate of the triadic representational relation. Moreover, although Interpretants are embodied in the interpreter/receiver of signs, they do not exactly correspond with the human agent, conceived as an individual subject that consciously interprets sings. The human agent is just the material vehicle of Interpretants, I would say. She functions as a *constraint* for the emergence of Interpretants: since those ones are feelings, reactions, generalizable actions, they require that somebody, with that body, that history of experience and so on, embodies them. By conceiving of the notion of Interpretant in this way, namely by stating the *logical primacy* of the agency of Interpretants on the subject-interpreter's individual interpretative acts, Peirce avoids the dyadic relation subject-object the classical notion of representation is built on. There is not an object-representation interpreted by a subject, but there are Interpretants caused within the boundaries of the representational relation. This means that interpretative acts do not take representations as the objects of their cognitive manipulation. On the contrary, the sign-vehicle determines its interpretation (and, according to Peirce, this makes it active; MS 793). This interpretation, on its turn, causes some effect on the second correlate and on the way it is connected to the first correlate: meaning acquisition takes place in the triadic mediation.

On the other hand, the Interpretant itself is thought to be *active*. The interpretation of representations is no longer thought to be something that has the "mental mode of being" (CP 5.473), if "mental" refers to an internal state,

such as a mental representation, but it is conceived as something that, passing by emotional and bodily effects, leads to the acquisition of a new habit of action. This last one, the habit of thought (which is the highest degree of knowledge, since it endows the subject with adaptive abilities in her environment), has nothing in common with the crystallized propositional knowledge the classical notion of interpretability implicitly refers to (i.e. acts of interpretation conceived as knowing *that* this representation is a token of that representation type). On the contrary, it has a *practical nature*. Indeed habit is a "general *operating* rule within the organism" (W4: 249; italics added). Its nature is operative, not intellectual, since habit is a "readiness to act in a certain way under given circumstances and when actuated by a given motive" (CP 5.480). Habit is as an action-schema, it is the *context-sensitive*[72] *knowing-how* to bring forth the practical bearings (MS 318: 22) of the object, which is the second correlate of the triadic relation. To put in other words, interpretation is the ability to act out cognition afforded by a representation.

This approach to representations gives very different insights from those offered by representation-friendly cognitive sciences. At this point of my explanation, I do not consider the broader implications of Peirce's semiotics for cognitive sciences (to some extent, this job will be done in §II.6). What I want to suggest here is that a comparison between the notion of representation in cognitive sciences and the way Peircean semiotics conceives of signs can be helpful to reassess the representational criteria endorsed by cognitive sciences. In particular, by taking into account Peirce's triadic definition of sign, it can be said that:

i) the interpretability condition does not entail the passivity condition necessarily (on the contrary, in Peirce's semiotics it implies activity);
ii) the internality condition, that is the idea that representations lay inside of human heads, is just an arbitrary choice made by cognitive scientists. Indeed, in the classical account of representation offered by Peirce, representations involve media distributed among the agent and the environment (for example, the musical notes, the perceived sounds, and the agent's feelings).

[72] The aspect of context-sensitiveness is explicitly expressed in the quotation above, where Peirce talks about the activation of habit-occurrences, and in the famous essay "How to make our ideas clear" as well, where he states that perception causes us to act (W3: 265).

This leads me to think of a reduction of the six criteria endorsed by cognitive scientists to four conditions: i) genuine duration; ii) decouplability; iii) standing-for; iv) interpretability.

II.5 Objections against Clark's AORs: standing-for, Interpretability, Decouplability and the problem of activity in representation. Some Peircean suggestions.

In the previous paragraph, I discussed the criteria philosophers of cognitive sciences endorse to define representations, and I reassessed them, explaining why I think that some of them are necessary while others are not. The outcome of this discussion has been a reduction of the representational criteria to four conditions: i) genuine duration; ii) decouplability; iii) standing-for; iv) interpretability.

To check if Clark's action-oriented representations meet these criteria, and then to determine if they are genuine representations or not, consider again the way he conceives of these representations.

Action-oriented representations are neural states triggered by a given aspect of the environment perceived by an agent engaged in an on-line cognitive task. Their content is both descriptive and prescriptive: they represent a state of affairs in the agent's environment, by encoding it as the content of an action-command. To give a propositional description of this action command, it can be expressed in the following way: "there is X, perform the action Y upon it".

It can be said that those representations

i) have a genuine *duration*: they are neural states activated in the context of a task, at a given moment in time.

ii) *Decouplability*: those representations do not meet the criteria of strong decouplability (and Clark explicitly says that; Clark 1997a: 144). Those representations are said to be triggered by a given aspect of the environment perceived by the cognitive agent. In the absence of this perceived aspect AORs are not tokened. In that case, other representations would be activated in the brain, but not those which represent what is perceived as an affordance for that individual action. Moreover, the indexical aspect Clark ascribes to his AORs is not consistent with the very notion of decouplability, to the idea of a representation that can be used-off line. Indeed an indexical representation is a representation that entertains a relation of "correspondence in fact" (W2:56) with what it represents. This is to say that we talk about indexes properly when there is a "direct physical connection" (W5: 254) between the index and the item represented. Then, the semiotic features of Clark's AORs make them tied to an existent, present context.

Nevertheless, if one would like to defend Clark's AORs, she could try to see a "little bit of decouplability" by thinking about the simulative aspect of AORs. In paragraph II.3, I said that AORs simulate what should be done by the agent in the environmental space as a response to perceptual stimuli. Given that the action command is part of the content of action-oriented representations, and given that the action that involves those perceptions has not been performed already when the action-oriented representation occurs, one could think that this conception of AORs, to some extent, entails decouplability. The representation stands for something that, when that representation is tokened, is not present in the environment, then it is not in causal contact with its immediate environment. Nevertheless, I think that this way to think about decouplability is too loose. Decouplability refers to the possibility of an item X to occur in a cognitive process without being in causal contact with the "current context" (Gallagher 2008a: 351). In the cases taken into account, "context" means a) the physical environment where the task takes place, b) the temporality of the task. Since AORs are said to be part of the dynamics of an action-perception loop, and since actions take time to develop, the fact that AORs represent something that is not present in the environment when they are tokened does not imply that they are decoupled. If it is true that, when they occur, that particular motor movement prescribed by the representation has not been performed already, it is also true that the action the future bodily movement is part of is already there, it has already begun. Indeed the agent did something (i.e. she moved her body) in order to give a rough structure to her perceptual array. Then Clark's AORs are not decoupleable from the spatio-temporal context of the action they are supposed to explain, not even in the loose sense of "decoupleable".

iii) *Standing-for/interpretability*: as said previously, in many passages Clark refers to AORs as "neural populations". The question is whether a neural population *in itself* can be considered to be genuinely representational. Usually cognitive scientists claim that neural structures are representations because they respond in a reliable manner to certain conditions. This is to say, since neuronal states lawfully (namely regularly) co-vary with perceptual stimuli, they are said to represent/carry information about what caused those stimuli, namely an object or feature of the world. What is not clear is how this correlation between a neural state and specific stimuli entails *per se* (and not according to the scientists' second order gaze,

which interprets neural states as representations) a *mediation* (Ramsey 2007: 120-121).

There are many reasons why the use of a representational vocabulary to denote the relation between neural states and perceptual stimuli is not clear. For example, Hutto and Myin (Hutto, Myin 2013: 63-71, and elsewhere again Hutto; Hutto 2011, 2013) make an objection to this account of representation by focusing on the problem of content. Representations are contents bearers. Carrying a content is a condition that, according to cognitive scientists, an item should fulfill in order to be called "representation". To talk about content there must be a) specified conditions of satisfaction to determine whether or not the content an item or process X bears is an accurate or true content; b) this entails that there is some X that can potentially understand if those conditions are fulfilled or not. In the account of neural activities as representational this "understanding X" is missing, at least if we consider what supporters of AORs write in a literal way. Indeed, as pointed out previously, AORs are usually thought to be dyadic representations: the interpretative function is not part of them.

Also one of Gallagher's objections against AORs, even if it is not focused on the problem of content, looks at the *missing third* in order to demonstrate that AORs are not genuine representations. As explained in the previous paragraph, representations are relations of mediation: an item X, by standing-for the item Y, carries information I about Y. What connects the item X with the item Y is an interpretative function. Gallagher, by recalling the interpretation of Peirce's concept of representation that Menary gives in his book *Cognitive Integration* (Menary 2007: 95-102), claims that neural populations cannot be considered to be representations because they do not have this triadic structure. On the one hand, he claims that the experiencing subject is not an interpreter of her brain states (then here he conceives of the Intepretant as the human interpreter, restricting the more loose concept of consumer that Menary uses[73]). On the other hand he says that neither the brain itself is an interpreter of its owns states, "unless one is willing to say that one process in the brain interprets another process in the brain as a sign

[73] Indeed Menary talks about genuine representations as items whose salient features are exploited by some consumer X, and he conceives of these elements of the representational relation as "established by conscious intention in humans, or by non-conscious and non-teleological biological function" (Menary 2007: 99).

of something happening in the environment" (Gallagher 2008a: 361).

Broadly speaking, I agree with Gallagher. Where there is not a triadic relation, there is not representation, because an item X, in order to stand for the item Y, should be interpretable. The point I would like to consider more closely in order to understand if it rules out the concept of representation in relation to brain states concerns the Interpretant as a human interpreter. As explained in the previous paragraph, according to Peirce the Interpretant is not an equivalent of the human interpreter. The human subject *embodies* Intepretants, but she is not the Interpretant itself.[74] As already mentioned, according to Peirce there are emotional Interpretants (feelings, emotions), energetic Interpretants (reactions, efforts), and ultimate logical Interpretants (habit changes). Then, if one aims to be faithful to Peirce's semiotics, she should not claim that, in principle, there should be a deliberate act of interpretation to talk about representational dynamics properly.

Nevertheless, it seems to me that Gallagher's objection can still be useful because it points to the right direction. As explained in the Peircean example of musical experience, the triad of Interpretants can be interpreted in terms of cognitive degrees. The feeling or emotion is followed by a reaction or an effort, which, by means of experience can become general, changing the experiential habits the subject is engaged with her world by. My question is: can brain states be interpreted as feelings, and then as the "first" Interpretant of a chain of Interpretants? If we follow Peirce's discussion about the modification of Interpretants within the temporality of semiosis, this not seems to be the case. Indeed, what follows the emotional Interpretant is an energetic Interpretant, which focuses the subject's attention on what she *felt*. Then, feelings, in this perspective, entail a quality the human subject has somehow access to[75]. Therefore,

[74] This point not only can be explained by recalling the triad of Interpretants mentioned here, but it is also made clear by Peirce himself. For example, in a letter to Lady Welby written in 1908, he says that to describe signs as determining an effect (an Interpretant) upon a person, is a "sop to Cerberus" (SS: 80-81), and that his semiotic conception concerns the broader activity of signs.

[75] This interpretation is motivated not only by the fact that Peirce says that the emotional Interpretant is followed by an energetic Interpretant (that is a mental or physical effort), but it is also faithful to the passages where Peirce talks about feeling more in general. For example, he writes that "We are immediately aware [...] of our present feelings" (CP 1.167), he talks about feelings experienced in outwards sensations" (CP 1.308), he speaks about the "vividness of a consciousness of the feeling" (CP 1.309), and about feelings as "immediate to consciousness" (CP 1.310).

> although Peirce's explanation does not entail that Interepretants are human Interpreters, it supports Gallagher's point: there must be a kind of "understanding" (even if this does not entail a propositional or complete understanding) of the sign-vehicle in order to give meaning to it. In this sense, Peirce's explanation also rules out the second option to argue in favor of AORs, namely the idea of considering processes in the brain as interpreting each other as signs of things or states of affairs in the environment. Brain states can be seen by scientists *as if* they were representations, but, to the extent that there is not actually one point within the flow of brain states that corresponds to a feeling, an effort or a disposition to act, they cannot be considered to be Interpretants of sign-vehicles by themselves.
> This "missing interpretability" is a crucial point, for at least two reasons. First, this rules out the possibility of an item X to bear a content: according to explanations in philosophy of mind and cognitive sciences, this makes an item non-representational. Second, "missing interpretability" undermines the first representational-condition: the ability of the item or process X to stand for the object or event Y. Indeed a sign-vehicle cannot stand for Y without an "understanding" that sees it as able to stand-for Y.

This exam of the representational criteria applied to Clark's AORs gives reasons to endorse hypothesis ii) about representations in action-perception loops formulated in II.3. There I said that I was *prone* to think that we do not need AORs to explain perception. In virtue of the discussion developed here, I *claim* that the use of a representational vocabulary to explain the cases taken into account her is unwarranted. The mechanisms Clark refers to when he talks about action-oriented representations are not representations because they do not meet three of the four criteria necessary to define an item or process X as a representation.

What I would like to take into account now is another problem that concerns the use of the concept of AOR in the Extended Mind Hypothesis. The objections to this approach considered till this moment have been developed at a formal level of explanation. The argumentative structure of those objections was the following one: action-perception loops are not representational because what is called action-oriented representation is not a representation.

I think that another way to consider the issue of AORs is to look at the role they play in the philosophical project of extending the mind. Do AORs play any interesting role in the philosophical job of "extending the mind"?

At first glance, it seems that the very idea of action-oriented representation is not consistent with an extensive account of the mind. This because the use of this epistemic posit transposes at another theoretical level the problem taken into in account in §I.9, namely the weak idea that the mind is "extended", that is the idea that there is something like an internal core starting from which the mind blooms towards the outside. The way Clark deals with AORs expresses this point hidden in the folds of the Extended Mind Hypothesis: the transcendental condition of action-perception loops is the activation of something embedded in human heads, a control and directive structure that causes action, which, in its turn, influences perception. The mind extends into the body and the world -namely the agent performs action-perception based situated cognitive tasks- thanks to the job of peculiar kinds of internal structures, which allow the agents to display a given kind of skill, and to act in a given environmental context adequately. This is to say that not only the EMH general ideas entail an internalistic commitment that, I claim, undermines the project of extending the mind, but also the individual explanations of certain kinds of cognitive processes (i.e. selective responsiveness to environmental stimuli) presuppose an epistemic primacy of intracranial resources on the external ones.

This coarse-grained idea about the way action-oriented representations do not match the requirements of an extensive and location-impartial theory of cognition and the mind can be also supported by fine-grained objections about AORs and mind-extension. As already said in the first paragraph of this chapter, in many passages of his work, Clark suggests that mind-extension is dependent on an active and dynamical account of cognition.

Then the question I think we should work on in order to understand the role played by AORs in the project of mind-extension is this one: does the way Clark describes AORs actually account for *representational dynamics in action*? To put in other words, do Clark's AORs account for action in representation?[76] Are AORs suitable to explain cognition, and then the mind,

[76] The way I consider this point, by talking about "action in representation", is an implicit reference to Mark Rowlands's discussion offered in *Body language. Representation in Action* (Rowlands 2006a), aimed to provide a representational account of action. Broadly speaking, I think that the way he deals with the problem of the relation between action and representation, if we consider the broader debate on AORs, is partial. Indeed, in order to argue in favor of action-oriented representations in cognition, and to undress them from passivity, one should first look at very activity of representations in themselves. This because the project of AORs

as a continuous process that crosses the bounds of the skull? I claim that AORs do not fulfill this job.

To explain this point, I begin my discussion by considering again what AORs are according to Clark. AORs are neural states which are said to bear a content that is at the same time descriptive of features of objects and prescriptive of motor commands. Because those representations are goal-directed or action-oriented, they are said to be active. The idea AORs are built on is that

> "the brain should not be seen as primarily a locus of inner descriptions of external states of affairs; rather, it should be seen as a locus of inner structures that *act as operators* upon the world via their role in determining actions" (Clark 1997a: 47, italics added).

Is this enough to claim that there is action in representation? I think that it is not, because the way action is involved in representation is too weak and loose. Indeed, a command, an invitation for action is just an *antecedent of action*, not an activity in itself. To account for action *in* representation, action should be considered to be one of the correlates of the representational dynamics: the dyadic conception of representation Clark's account of AORs is built on does not allow action to be part of the representation itself. Explaining the way Clark describes AORs by making use of the semiotic terminology previously presented, it could be said that AORs of *Being there* are the relation between a sign-vehicle, that is a neural population, and an object (the item of the external world represented under a certain action-respect, for instance as graspable, drinkable, and so on). Action, namely what the agent actually does upon the external world, falls outside of the representation. It is not part of it, but it is rather conceived as an outcome of internal representational dynamics. This explanation, instead of getting rid of the cognitivists' dichotomies such as action/cognition, inside/outside, representational structure/world, endorses them. More precisely, even if the EMH, at a general level, claims that action is part of a cognitive process, and then that cognition is *in* action, when it deals with AORs, it seems to consider action just as a "derivative" mode of cognition, I would say. Action, even if it is cognitive, is an output[77] of an internal representational mechanism that is not active by itself.

---

was first motivated by the necessity to give an active account of cognition and the mind which still involves representations.

[77] Indeed Clark says that action is *determined* by internal representations (Clark 1997a: 47).

Moreover, it is not clear how Clark's AORs point to a useful theoretical direction to account for the dynamical and processual features of cognition (Clark 2008: 24), namely to the idea that cognition is realized by continuous interactions between internal and external structures. Indeed, Action-oriented representations are neural *states* which cause the cognitive agent's behavior: they seem to be just a repetition of the idea, critically considered in §I.9, that the mind is a container of mental states, connected to action and embodied dynamics just by a causal or functional relation. To explain this point better, I would say that, to account for representational dynamics, for the continuity between acting, perceiving and representing, the theory should to think of a shift from representations as mental *items* or *states* to the *activity of representing*. Indeed, by thinking of representations as neural populations, Clark seems to end up with reifying representational mechanism: action-oriented representations are those "things" inside of human brains that cause the agent to act in a certain way. The real dynamic and kinetic aspect of cognition (acting in a certain way, moving the body, acquiring a certain bodily posture, and so on) is out of the representation. The fact that action is said to be caused or controlled by internal representations does not make it continuous with the representational mechanisms, and this, again, because action is not part of the discrete items that Clark calls "representations". To make representations active and dynamical, they should no longer be conceived as things or states, but i) as processes by means of which cognition is brought forth, and ii) as activities.

A possible way to think of the activity of representing -explaining representations not as items or states but rather as active processes that break the internal/external divide, or better, that have an "extra-cranial foundation"- can be found again in Peirce's semiotics.

The theoretical foundations of Peirce's semiotic account of cognition has been sketched out already. Nevertheless, to explain the theoretical value of Peirce's semiotics for the current debate in cognitive sciences better, it is worth to make the theoretical consequences of Peirce's semiotics more explicit.

Peirce uses the word "semiosis" to refer to any

> "action, or influence, which is, or involves, a cooperation of three subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs. [It is] the action of almost any kind of sign" (CP 5.484, text into brackets added)

"Semiosis" is a word that Peirce uses to refer to signs *in actu.* The suffix of the word points to this active aspect, to the idea that semiosis concerns the activity of signs. Indeed, Peirce borrows the word "semiosis" from the Greek (in particular, he takes the expression from Philodemus; Fisch 1986: 241; 274) "semeiôsis". The Greek suffix *–sis* exactly expresses the ideas of activity and process: it means "the act, action, activity or process of" (Colapietro 1989: 19).

The continuous and processual feature of semiosis suggested by the etymology of the word and explicitly ascribed to sign-action, is a consequence of some epistemic considerations clearly expressed in the so called anti-Cartesian essays. In particular, in the essay "Questions concerning certain qualities claimed for man" and in "Some consequences of four incapacities", both published by the *Journal of Speculative Philosophy* in 1868, Peirce develops an anti-intuitionist, anti-introspectionist (Fusaroli, Paolucci 2011: 17-18), and anti-psychologist (Calcaterra 2008) theory of cognition according to which:

i) Cognitive agents neither have an introspective access to their knowledge of X, nor they have the power of introspectively distinguish the different items that are part of a set of knowledge[78].

ii) Cognitive agents do not have an immediate cognitive access to external facts or objects, but their access is mediated by some kind of previous, grounding knowledge (W2: 209-211).

iii) Actual thoughts (namely occurrent thoughts, which Peirce defines as feelings) do not have an intellectual value, because they are not general. To acquire meaning, thought should be connected to subsequent thought by means of a sign, which makes knowledge increasing in virtue of the mediation of an Interpretant, in an ever receding process in which the ultimate meaning of a thought shows up as virtual, as something that will be done or achieved in the future (W2: 227).

iv) Then, cognition is a semiotic "continuous process" (W2: 224), namely a continuous relation among triadic relations whose Intepretants become the first correlate of new semiotic relations. Thanks to these semiotic relations, something new concerning an object, a fact, or an event shows up in the "train of thought" (W2: 224), by taking with it the history of this novelty (knowledge it belongs from), and by affording new cognitive acts.

[78] For the argumentation against Descartes and the Cartesian lineage see W2: 193: 200; 207.

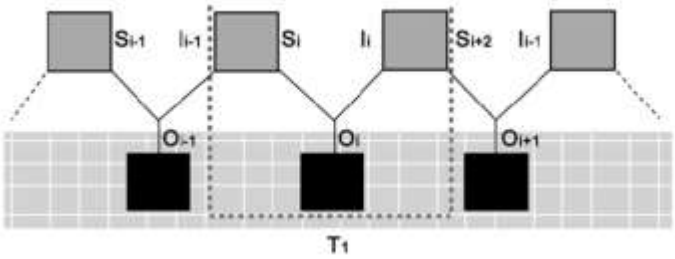


Figure 4: Chains of Triads [Picture taken from Atã, Queiroz 2016]

The semiotic approach to cognition is then, according to the Peircean definition, dynamic and processual: cognition emerges by means of signs, and signs are part of a process, the *interpretative* one, which makes them significant, meaningful (Deacon 2014: 97).

Also, semiosis, namely sing-action, is in principle unlimited (MS 599: 28-36). Even if a final Interpretant can be said to stop a chain of signs (namely what Peirce calls "the train of thought"; W2: 224), and this could make us prone to think that semiosis at that moment is in a certain *state*, this stop should not be interpreted in a chronological sense (Eco 1976: 1465), but rather in a logical sense. The Final Interpretant, which is a habit-change that modifies and enhances the cognitive agents' experience, is said to be "final" in a teleological sense. The

> "Final Interpretant is [...] the effect the Sign *would* produce upon any mind upon which the circumstances should permit it to work out its full effect. [The] Final Interpretant is the one Interpretative result to which every Interpreter is destined to come if the Sign is sufficiently considered. [...] The Final Interpretant is that toward which the actual tends." (SS: 110-111, Peirce's letter to Lady Welby, 1909; text into brackets added)

The final Interpretant, which brings into light the whole conception of an object, is the habit-change semiotic chains *tend to* and that would modify everybody's experience[79] of that object in the future (CP 2.148; see also Nöth 2016: 39). This is to say that a clear cut, identifiable with a state in the semiotic process, is just envisioned: it is a virtual stop. It is a tendency, a final task, a "hoped" achievement, not an actual state of the semiotic process. Moreover, the *telos* semiotic chains tend to is not conceived as a state, but it is itself active,

[79] According to the pragmatist approach to cognition, experience should be understood as the actions that can be performed upon and with that object (W3: 266 and ff.).

because the final Interpretant is identified with a peculiar kind of practical knowledge, namely with knowing-how to act in certain circumstances (habit).

To put it in other words, if semiosis is taken into account at a macro-level of analysis, in its whole development, it can be said to unfold through the *continuous* and reciprocal *relation* among sings. In this process, which entails the growth of knowledge, individual components and agents are not actually conceived as individual, discrete units, but they are rather described as *transitions:* signs, in their being influenced by previous signs and in their tending towards other, more developed signs (Interpretants), are "always about to start" and "never about to end" (Fabbrichesi, Leoni 2005: 39).

To explain this point better, as Fabbrichesi suggests in her book *Continuità e variazione* (Fabbrichesi 2005: 39-44), it is worth to take into account an example given by Peirce in his essay "Questions concerning certain faculties claimed for man" (1868; W2: 193-210).

> "Suppose an inverted triangle ▼ to be gradually dipped into water. At any date or instant, the surface of the water makes a horizontal line across that triangle. This line represents a cognition. At a subsequent date, there is a sectional line so made, higher upon the triangle. This represents another cognition of the same object determined by the former [...]. The apex of the triangle represents the object external to the mind which determines both these cognitions. The state of the triangle before it reaches the water, represents a state of cognition which contains nothing which determines these subsequent cognitions. To say, then, that if there be a state of cognition by which all subsequent cognitions of a certain object are not determined, there must subsequently be some cognition of that object not determined by previous cognitions of the same object, is to say that when that triangle is dipped into the water there must be a sectional line made by the surface of the water lower than which no surface line had been made in that way. But draw the horizontal line where you will, as many horizontal lines as you please can be assigned at finite distances below it and below one another. For any such section is at some distance above the apex, otherwise it is not a line. Let this distance be a. Then there have been similar sections at the distances 1/2a, 1/4a, 1/8a, 1/16a, above the apex, and so on as far as you please. So that it is not true that there must be a first. [...] Say that instants and lines are fictions; only say, also, that *states of cognition* and judgments are *fictions*. The point here insisted on is not this or that logical solution of the difficulty, but merely that cognition arises by a process of beginning, as any other change comes to pass" (W2: 210, emphasis added)

As the quotation above shows, according to Peirce cognition cannot be grasped properly by the concept of state: if the word "state" occurs, it should be understood in a fictional sense only. "State of cognition" is a metaphorical phrase used by philosophers to understand semiotic processes, focusing on some of their phases. Cognition can be compared to the lines made by water on the surface of a hypothetical triangle dipped into water. The traces left by water on the surface of the triangle form a continuum of traces that goes from the basis of the triangle towards its apex, which stands for the object that has to be known. Every time someone tries to identify the borders of each line, she finds herself engaged in the difficult (or impossible, Peirce would say) task of moving the beginning and the ending points of that line, because that line keeps on transiting towards previous and subsequent lines. The bounds of each piece of cognition cannot be identified *de facto*, because what should be found keeps on changing, it continuously flows into what, according to a second order gaze aimed at schematizing those continuous transactions[80], can be said to come "before" or "after" that piece of cognition. To put it in semiotic words, since according to Peirce signs are the ingredients of cognition, cognition is made of the continuous process through which signs of the "semiotic cable" pass the one into the others, keeping on being so intimately connected that they cannot be distinguished as discrete elements (W2: 213 for the use of the metaphor of the "semiotic cable).

To this conception of the semiotic activity, which accounts for the processual/relational ontology of it (Parker 1994; Pape 2014; Atã, Queiroz 2016), namely for what actually happens with signs, and for their constitution as well, Peirce adds a *methodological* level of explanation, which is the one taken into account in the previous paragraph, namely the description of the *micro-semiotic level* (Atã, Queiroz 2016: 115 and followings) in which the phases of the semiotic process are abstracted from the process. Acknowledged that semiosis, according to the law of continuity or *synechism*, is a continuous process whose elements or parts are not actually distinguishable as individual items, the philosopher can give a *formal* description of signs, looking at the relations that take place in the triads S-O-I (Sign-Object-Interpretant) distributed in the semiotic process.

[80] This is a very rough description of the cognitive and metaphysical principle of continuity or *synechism* (from the Greek *synechismos*, continuous; EP 2:1), which is the "doctrine that all that exists [for example mind and matter, or thought in consciousness] is continuous" (CP 1.172; text into brackets added. See also the paper "The Law of Mind", published by the *Monist* in 1891; W8: 135-157).

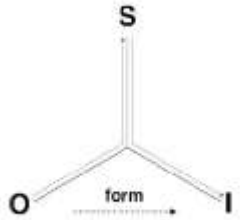


Figure 5: Micro-level semiosis
[Picture taken from Atã, Queiroz 2016]

As already explained -and as Figure 5 shows- a sign, namely a representation, is a triadic relation that involves a sign-vehicle, an Object, and an Interpretant. Looking at the internal relation that constitutes a sign, it can be pointed out how Peirce's semiotics tries to account for action in representation, or better, of signs as the activity of representing. In fact the third element of the semiotic relation, the Interpretant, especially if it is conceived as an ultimate Intepretant, as the higher stage of the interpretative development, is said to be an *action* or an *action-schema* (habit). The crucial point of this description of signs consists again in the triadic constitution of representations. By conceiving of the interpretative function of signs-vehicles as part of the sign itself, and by defining the Interpretant as an action or a pattern of action, Peirce "activates representations". Or, better, since the Interpretant is the crucial part of the semiotic process -indeed it is what allows interpretability, and interpretability, according the representational criteria previously given, is what makes an X a representation- this explanation leads to the collapse of the concept of representation on the *activity of representing*. Indeed, the Interpretant is a representation too: it is a representation that mediates between the first and the second correlate, and it represent their relation by *acting upon the external world.*

To consider this point better, think again of Clark's concept of action-oriented representation seen from the lens of Peirce's semiotic terminology[81]. In the example given in §II.3 we saw a cognitive agent who was performing the task of picking up all the objects designed for drinking. According to Clark's

[81] It is worth to make clear that in my development of this reworked notion of "active representation" I do not take into account the overall project of Peirce's theory of perception, neither I am faithful to his inferential account of perception, whose core idea is "perceptual judgement" (MS [3] 881, *Telepathy*, partially published in the *Collected Papers*; CP 7.597-688). This because the aim of my research project does not consist in a systematic comparison between Peirce's philosophy of cognition and contemporary approaches to cognitive sciences. My aim is rather more modest, and it consists in showing how some specific concepts of Peirce's philosophy and semiotics can be helpful to understand and to enrich some problems of the current debate in cognitive sciences.

explanation, the action of picking-up, which was conceived as an outcome of an internal representational mechanism, was said to be performed thanks to the occurrence of action-oriented representations that have this form: "there/object/for/drinking/-grab/it!". Considering Peirce's semiotics -a least at a provisional and general level[82]- it could be said that the relation between action and perception that Clark explains by means of the concept of action-oriented representation can be modeled in the following way. There is a representation whose correlates are stimuli from the environment (first correlate), an object of the environment, namely the object designed for drinking (second correlate), and an Interpretant (third correlate), that is the action of grabbing the object for drinking. The action of grabbing the object designed for drinking shows (and, according to Peirce this means "represents") the "affordance power" of the first correlate: the Intepretant actualizes a possibility of action by performing that action. At the same time, the Interpretant, by means of the mediation of the sign-vehicle, communicates the form of the object represented (MS 793: 1; EP 2: 544; Atã, Queiroz 2016: 110). By making the agent acting in a certain way upon the object, it explicates (namely it "spreads out", "unfolds" from the Latin *ex,* out, and *–plicāre*, to fold) the meaning embedded in the object, which the pragmatist tradition identifies with its practical bearings[83], and with the form of action that actualizes them.

In this sense, by appealing to the general idea of sign in Peirce, it can be said that semiotic representations can be considered to be "an *improved* version" of Clark's AORs, and this at least for two reasons. First, not only they are endowed with a conative power (because the first correlate of the representation affords something that can be done, according to the form of the object represented), but, thanks to the concept of Interpretant, they are also endowed with a genuine active ability. To put it in other words, Peircean signs account for the antecedent of action, for the object the action is performed upon, and for action too. Second, those representations do not entail the internalistic prejudice Clark's AORs suffer from. In Clark's explanation AORs were internal action-controllers. In Peirce's semiotics, thanks to its entanglement with pragmatism, representations work as controllers of action as well, but they do so by operating "out of the head"[84]. Indeed, what constrains

[82] An attempt to apply Peirce's semiotics to the more specific problem of contextual relevance will be made in the next paragraph, where I will try to reassess the enactivist proposal by integrating it with Peirce's concept of indexicality.

[83] See the 1878 essay "How to make our ideas clear" (W3: 257-275).

[84] Notice that the externality of the mind in respect to the boundaries of human heads does not lead to conceive of the activity of signs as non-mental. Rather, by conceiving of cognition as the activity of signs constrained by objects of the world and regulated by habits of action, the

the action that will be performed is an object of the world[85], which is the second correlate of the semiotic relation (defined by Peirce as "immediate object"; CP 4.536). Moreover, the overall cognitive process that unfolds through cables of semiotic triads is regulated by something that is internal to the semiotic relation but that is not inside the human brain: habit conceived as a law-like rule of action, which explicates itself when it solicited by certain (actual, or conceivable, according to Peirce's late pragmatism, which he calls "pragmati*ci*sm"[86]) conditions.

Peircean explanation contributes to define the mind as an activity that takes place in the interaction between the cognitive agent and the environment, anticipating a theory of the "external mind" (Paolucci 2011: 78-86). In this theory, the mind is not thought to be dependent on brain structures (Peirce himself writes that "thought is not necessarily connected to a brain"; CP 4.551), but is rather explained as "an external sign" (CP 5.313-4) that develops by means of semiotic relations.

[85] See MS 793: 1; EP 2: 544; see also MS 499, where the object is said to be the sign's determinant", and MS 320:7, where the object is said to be "the cause of the sign".

[86] See for example the letter to Mario Calderoni written in 1905 (CP 8.208), "Issues of Pragmaticism", published by *The Monist* in 1905 (CP 5.438 and ff.), "A Neglected argument for the reality of God", published by *The Hibbert Journal* in 1908 (CP 6.490).

II.6 Enacting salience. Reassessing the enactivist proposal through Peirce's indexicality.

In the previous paragraph I considered some objections to AORs. My argument was twofold. First, I argued against AORs by developing my objections at a formal level. I showed that Clark's AORs do not fulfill three of the four representational criteria I gave in paragraph II.4 as an outcome of my comparative analysis of cognitive scientists' representational conditions and Peirce's concept of sign. Then, I claimed that AORs are not actually representations; therefore I concluded that if we buy this epistemic posit, we are committed to an unwarranted philosophical argumentation. Second, I developed my objections to Clark's AORs at another explanatory level: I showed that AORs do not even play an interesting job in the broader economy of the philosophical project of mind-extension. I concluded my discussion by suggesting that Peirce's concept of sign does not suffer from the problems Clark's AORs entail, which make them unable to support the full hypothesis of mind-extension. This discussion was meant to be a preparatory stage for what will be discussed in this paragraph.

Here I will take into account the problem AORs were supposed to solve form another, non-representational (at least in a "classical" sense) point of view. In fact, I will take into account how an explanation grounded in Merleaupontian phenomenology and in ecological psychology tries to solve the Generalized Frame Problem. This one has been defined as the problem of selecting (i.e. framing) only relevant information distributed in the context of a cognitive practice (Chow 2013: 312- 315), by keeping track of relevant changes in the environment (Haselager, Van Rappard 1998: 161).

I will suggest that this account of selective responsiveness to contextual relevance is an interesting response to the philosophical problem taken into account. This for at least three reasons. i) It does not build its explanation on an epistemic posit that is unwarranted; ii) it accounts for the agent's phenomenological experience (and this seems to fit with 4Es' broad project of accounting for everyday cognitive experience; see §I.2); iii) it is not built on an implicit internalistic prejudice.

Nevertheless, despite I endorse this "radical embodied" or "enactive"approach, I also claim that some aspects of the explanation can be developed further. If on the hand the fully embodied and anti-representational explanation, by focusing on the concepts of affordance and affect, gives a non-brain centered solution to the Generalized Relevance Problem, on the other hand it seems to be too tied to one of the poles of the agent-environment relation: the agential one. Hence I suggest that some insights useful to work on the role played by the environment in selective responses can be found in Peirce's thought. In particular, I suggest that Peirce's concept of index can be

helpful to explain why "some affordances stand out more than others" (Bruineberg, Rietveld 2014: 2), point at the core of the "radical embodied" explanation.

I begin my discussion of the non-representational account of the frame problem (Bruineberg, Rietveld 2014; Dreyfus 2002a; 200b; 2008; Rietveld 2008a; 2008b; 2012a; 2012b; Rietveld, Kiverstein 2014) with a very rough sum-up of some insights form ecological psychology. Indeed, many of the contemporary non-representational approaches are inspired by Gibson's theory of perception (Gibson 1979).

The core idea of Gibson's ecological approach to perception is that perceptual activity consists of the intentional movement of the whole being in its environment. Perception, rather than being a pre-requisite for action, is a mode of action.[87] What we perceive is not mentally represented, but is rather a *direct* function of how we act. Knowledge offered by perception is thus practical: "it is knowledge about what an environment offers for the pursuance of the action in which the perceiver is currently engaged" (Ingold 2000: 166). To put it in other words, human and non-human animals do not perceive neutral objects, but they directly perceive *affordances* for action[88]. What is in the world invites agents to act in a certain way and not in another one: environmental features directly trigger certain kinds of action, which are part of the flow of the practical engagement animals' lives unfold by.

According to Gibson,

> "The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes,* either for good or ill. [...] They are unique for that animal. They are not just abstract physical properties. They have unity relative to the posture and behavior of the animal being considered" (Gibson 1979/2005: 119-120).

Affordances are what the environment invites the (human or non-human) animal to do, according to its possibilities of action. In a rough sense, affordances can be described as fundamental features of things that determine how things can be possibly used by agents: "they provide strong clues for the

[87] For a broader picture of this active account of perception, see the more detailed explanation of "enactive perception" given in §II.2.

[88] Notice the similarity between this idea and Uexküll's *performance-qualities*, taken into account in §I.7, where I presented the concept of Umwelt. A more detailed discussion of the similarity between Uexküll's theory of Umwelt and Gibson's affordances can be found in O'Neill 2008, Chapter 9.

operations of things" (Norman 1988: 9). For example, a chair is for support, and then it affords sitting; a glass is for seeing through and for breaking; flat, porous, smooth surfaces are for writing on; plates are for pushing, knobs are for turning, slots are for insert things into, and so on (examples taken from Norman 1988: 9). And again,

> "an elongated object of moderate size and weight affords wielding, [...] a rigid object with a sharp dihedral angle, an edge, affords cutting and scraping, [...] a graspable rigid object of moderate size and weight affords throwing, [...] a hand- held tool of enormous importance is one that, when applied to a surface, leaves traces and thus affords *trace- making*" (adapted from Gibson 1979/2005: 125).

At first glance, this "being-for" of objects, which can be also said to constitute their "*action-orientedness*", seems to be determined by the very materiality of things: by the very material they are made of, by their shape, by the ergonomics and physiology of things, and so on. Nevertheless, even if affordances are sometimes defined as perceivable *resources* of the environment (Rietveld, Kiverstein 2014: 327), Gibson claims that they should not be identified with mere physical properties. This because they are thought in tandem with the animal's posture and behavior: affordances are invitations to act in a certain way given the *embodiment* of an agent.

To explain this point, Gibson, in his book *The ecological approach to visual perception,* offers a simple example (Gibson 1979/2005: 119-120)[89]. Consider a nearly horizontal terrestrial surface that is sufficiently extended (relatively to the size of the animal), and which is made of a rigid substance (relative to the weight of the animal). This surface affords support: for quadrupeds and bipeds, it permits an upright posture, therefore is "walk-on-able" and "run-over-able". The same thing cannot be said about a non-stable, fluid and liquid surface, like water or a swamp. For heavy terrestrial animals, instead of being "walk-on-able" or "run-over-able", the surface is "sink-into-able". On the contrary, this same surface affords support for insects, for example: "different layouts afford different behaviors for different animals, and different mechanical encounters" (Gibson 1979/2005: 120).

This suggests that affordances -as Chemero points out (Chemero 2003; Chemero 2009)- are intrinsically *relational*. Affordances are neither properties of the environment, neither properties of perceiving animals, but they are rather the relation between animals' abilities and perceivable features of the environment (Chemero 2003: 181). This is to say that, according to ecological

[89] For another example useful to understand the relation between perceived properties of the environment and the embodied agent, see my discussion of Uexküll's example of the rock (Uexküll 1982: 27), discussed in §I.7.

psychology, there is a direct link between perceived possibilities of action and the embodied agent's capacities (Rietveld 2008a: 341). Affordances are the relation between what an agent can do -and implicitly knows how to do, because she learnt how to behave in order to perform actions upon aspects of the world- and features of the environment that invite the agent to act in a certain way, fitting with the agent's embodied or "sensorimotor abilities" (Chemero 2009: 150).

Nevertheless, despite this tight connection between the animal's embodied abilities and resources of the environment that defines the concept of affordance, it should be noticed that affordances are not isolated, they are not environmental features that invite the agent to perform a *restricted* set of individual actions. On the contrary, according to Gibson, affordances are *holistically connected* and they are said to form ecological niches. Smaller affordances units are embedded in lager units, in which smaller units are nested. For example, "canyons are nested within mountains; trees are nested within canyons; leaves are nested within trees; and cells are nested within leaves" (Gibson 1979/2005: 5; 120). To put it in other words, animals are not simply exposed to the "climbability" of mountains and trees, to the "graspability" and "eatability" of leaves, but they rather perceive a complex intertwinement of invitations for action. This is to say that the animal, during its engagement with the environment, does not usually perceive singular or discrete affordances, but it is rather solicited by "nests of affordances". This nest of affordances determines *how* an animal lives in its environment (Gibson 1979/2005: 120), in its being solicited to perform complex flows of actions. This suggests that affordances can be defined as the relation between perceivable features of the environment and the animal's "way of life" (Gibson 1979/2005: 3; 120) or "form of life" (Bruineberg, Rietveld 2014: 2; Rietveld, Kiverstein 2014: 327-330), namely what is typical of that living being[90]. Moreover, it should be noticed that the complexity of ecological niches or "nests of affordances" is motivated by the very definition of affordance. As said above, affordances have a relational nature: they pertain to the animal's embodied abilities and to the possibility to perceptually exploit environmental resources. Abilities are not just mechanical dispositions that solicit an agent to act in a certain way in a given situation. Physical dispositions or reactions entertain a determined or deterministic causal relation with the situation that triggers them: when they are coupled with the right enabling conditions, dispositions are guaranteed to become manifest (for example, think about the soluble solid sugar that will always dissolve in water in suitable conditions;

[90] See Hunter's interpretation of Wittegenstein's *Lebensformen* in the *Philosophical Investigations* (Hunter 1968).

Chemero 2003: 189). A mere disposition "just is", it cannot be changed or improved. On the contrary, an ability is a practical know-how acquired through experience and training, its efficacy can change, it can be improved, and so on. This means that, during her practical engagement with the world, the agent can make her practical knowledge wider and wider, and this always according to her body, which constrains not only the actions the agent can perform, but also the abilities she can acquire. This possibility to acquire new embodied skills and to develop the abilities she has already acquired influences the structure of the "landscape of affordances" (Rietveld, Kiverstein 2014) the agent is engaged with. To the extent that affordances are perceived in virtue of the abilities the agent is endowed with, changes in the agent's practical knowledge (e.g. improvements, new acquisitions) are correlated with changes in the landscape of affordances practices take place in. An improvement or enrichment of the agent's practical knowledge determines an enrichment of the field of affordances: the agent endowed with brand knew abilities becomes sensitive and responsive to features of the environment that did not afford certain skilled action in the past.

For example, think about the main character of the thought experiment offered in §I.7: the climber. Before her training, mountains, walls, and buildings did not afford "climbability" in the agent's experience. Instead of being perceived as things whose "demand character" (Dreyfus, Kelly 2007: 52) consists in their being-climbable, those things used to invite the agent to perform other kinds of actions, for instance contemplating the natural beauty of the mountain or the fine architecture of the building, hanging up pictures to the wall, and so on. The acquisition of a new ability, that displays itself as an expertise, as a readiness to act enabled by anatomic and postural modifications (e.g. the arching shape of her hands, the strength of muscles and tendons, the relocation of the body's barycenter), modifies the agent's experience. Every time she sees a mountain, a wall or a building, the climber does not only deal with their beauty or with their possibility of being used as surfaces for hanging up, but she is also solicited by their climbability. Then, if this example is generalized, it can be said that the acquisition of new embodied skills can be interpreted not only as an improvement or enrichment of the agent's practical knowledge, but also as an "education of attention" (Ingold 2000: 354). Thanks to skill-acquisition, which takes place during embodied interactions with socio-material scaffoldings (Rietveld, Kiverstein 2014: 331; Ingold 2000: 354 and ff.), the salient environmental space (namely the field of affordances) is enriched. At the same time, the agent becomes ready to be responsive to this brand new salience.

This short exam of the notion of affordance shows that ecological psychology conceives of perception as a direct, non-representational activity, which involves an embodied agent with her sensorimotor abilities (and more in general her practical abilities) and a non-neutral, but rather action-oriented world. This action-oriented world can be said to be a rich "landscape" or "field" of affordances, in which each invitation for action is holistically connected to other invitations. Therefore, it can be said that the skilled agent is always situated in a world that asks her to perform a huge set of actions.

The question that seems natural to ask at this moment is the one I dealt with for the most part of this chapter: the Frame Problem issue. Given that fields of affordances are so rich and complex, how can we explain that agents are prone to being solicited by certain affordances of their practical field and not by others? To put in other words: why some affordances stand out more than others (Rietveld, Bruineberg 2014: 2)? Why agents usually cope with their environment just by acting adequately (Rietveld 2008a: 241), namely acting upon some environmental resources and not upon others?

To put it roughly, according to some enactive approaches to cognition, the phenomenon of skillful coping does not depend on a system of action-oriented representations or on a complex internal mechanism of action-planning and action-control, but it is rather a matter of attraction, or better, *situated attraction*. Attraction in a spatio-temporal context puts some affordances in the background, and turns other affordances into *solicitations* or affordances that are *relevant* (Rietveld, Kiverstein 2014: 342) in that specific situation.

To explain this "matter of attraction", the affordance-based explanations I am considering here recall Merleau-Ponty's phenomenology. As already pointed out, many passages of Merleau-Ponty's works (in particular his *Phenomenology of Perception;* Merleau-Ponty 1945/1962) anticipate the core ideas of the embodied and enactive approach to cognition (Dreyfus 2005 ; Varela, Thompson, Rosch 1991 ; Gallagher, Zahavi 2008; Gallagher 2008b; 2012; 2017). Here I will not take into account the whole story about Merlau-Ponty and recent cognitive sciences, but I rather focus on two important Merleaupontian concepts contemporary cognitive accounts focus on: *praktognosia* and affect.

By reassessing Husserl's notion of operative intentionality -according to which the agent's engagement with the world takes place through actions and projects non-reducible to mental states (Gallagher, Miyahara 2012: 120)- Merleau-Ponty claims that humans' access to the world is grounded on a particular kind of practical knowledge (*praktognosia*) that is bodily and motoric. Instead of being considered as a particular case of knowledge, logically and epistemically forerun by a symbolic or sense-giving function based on acts of

thoughts, embodied cognition is original, primary (Merleau-Ponty 1945/2002: 162-164). This is to say that agents' cognitive practices are primarily action-oriented and goal-directed. They are grounded on a kind of intentionality that is "determined by what the agent is doing and what the agent is ready to do [that is *praktognosia*]" (Gallagher, Miyahara 2012: 136, text into brackets added), and this kind of doing is performed by means of the body. To explain this point by recalling Merleau-Ponty's words it can be said that:

> "My body is the fabric into which all objects are woven, and it is, at least in relation to the perceived world, the general instrument of my 'comprehension' [...]. My body is the pivot of the world: I know that objects have several facets because I could make a tour of inspection of them, and in that sense I am conscious of the world through the medium of my body" (Merleau-Ponty 1945/2002: 273; 94-95)

This is to say that agents have access to the world thanks to that general instrument of cognition that is the moving and active body, endowed with what Merleau-Ponty calls "motor power", "motor project", or "motor intentionality" (Merleau-Ponty 1945/2002: 127). In cognitive practices, instead of thinking that the world is like this or like that, we implicitly know what *we can do* (Merleau-Ponty 1945/2002: 127), we are "motorically directed" towards the world, which is perceived according to our goals.

For example, think about a girl who is playing football. When the agent is performing this activity, to her, the football field is not an object of knowledge. She does not speculate about its qualities, but she just acts, she is smoothly engaged in the flow of her activity. The football field is not just made of grass, it is not just rectangular, and so on, but it is pervaded by lines of forces (e.g. the yard lines), and it is articulated in sectors (e.g. the openings between the adversaries) that call for a certain mode of action and which guide action as if the player were unaware of it. The field is not given to the agent as an object of knowledge, but it is rather "the immanent term" of "her practical intensions".[91] Merleau-Ponty's motor intentionality is precisely this: the football player's becoming one with the practical bearings of the field and her becoming one with her bodily posture. It is the relation between the goal-directedness of the agent's practices and the capacity of the world to afford the agent's actions.

Crucially, this pre-reflective "I can", that is *praktognosia*, is dynamic: the whole organism's task is redefined constantly, according to the spatio-temporal situation of the task. The leading 'I can' changes from moment to moment (Rietveld 2008a: 342). What the body tends to is always and crucially

[91] Example taken from Merleau-Ponty's *The Structure of Behavior* (1942), as reported in Thompson 2007: 313-314.

determined in tandem with what the agent perceives and with the momentary embodied condition of the agent. In this sense, the practical knowledge Merleau-Ponty talks about is highly context-sensitive, and this flexibility is precisely what the problem of selective responsiveness is concerned with.

Nevertheless, even if what Merleau-Ponty says about the fundamental 'I can' seems to support the picture of affordances given above, because it takes into account cognitive agents' motor abilities, the goal-directedness of the perceived world, and the relation between these two *relata*, it is still not clear *how* selective responses occur. To put in other words: what does direct a motoric intentional movement to that specific feature of the environment the adequate action will be performed upon?

According to contemporary enactive accounts of selective responsiveness, a possible solution to this issue can be found again in Merleau-Ponty, and in his way of characterizing motor intentionality as *affectively tuned*. As Rietveld points out (Rietveld 2008a: 345-346), Merleau-Ponty is quite clear about this point. Indeed, in *Phenomenology of Perception* he writes that

> "The perceiving subject ceases to be an 'acosmic' thinking subject, and action, feeling and will remain to be explored as original ways of positing an object, since 'an object looks attractive or repulsive *before* it looks black or blue, circular or square'" (Merleau-Ponty 1945/2002: 28, emphasis added)

Merleau-Ponty says that the way agents have access to the world is not only grounded on a fundamental action-orientedness, but also, in its turn, this action-orientedness is affectively or emotionally determined (Heidegger already noticed; see §II.2). Before being triggered by perceivable features of the world, the agent *feels* attraction or repulsion for those features, according to her own concerns. In this sense, the body, described as a "general instrument of comprehension", not only is a skillful body, which, according to its abilities, "is designed" to cope with the environment, but it is also a "concernful" body (Rietveld 2008a). It cares about what happens, and thanks to this affective mode, it is more or less opened to certain aspects of the world. That is why, given the embodied agents' motor abilities, some affordances leave her cold, and others attract her. The agent is always and already emotionally directed towards some features of the environment and not towards others because what she perceives first is an *affective situation.* This affective situation is constituted by her own concerns and affective states, and of the affective tone of the perceived environment, which displays attractive or repulsive features. To put it in other words, in situated cognitive practices some affordances stand out more than others because they matter more, and they do so because they are perceived as attractive. In this sense, it can be said that fields of affordances are

affectively structured, and they show up in this way because the perceiving agent's behavior is controlled at an affective level before being controlled at the level of action. A given affective tone is responsible of the initiation of a cognitive act (and then of the agent's being solicited by certain affordances) and for the bifurcation of two cognitive acts. This seems to suggest that agents act in an *adequate* way in a context because their perceptual activities are *affectively controlled*.

To explain this point better, Rietveld (Rietveld 2008a: 350-351) gives an example taken from Varela's article "The specious present: a neurophenomenology of time consciousness" (Varela 1999: 298-299). Consider an agent who is writing a paper on her laptop. While writing, she hits the control key, and she is shown a message that says "Do you really wish to erase this text?". She find herself in an affective tone of tension, which is followed by the awareness of the possibility of making a fatal mistake: if she clicks on the "yes" button, what she wrote will be deleted forever. The message triggered her attention, and this made the unreflective cognitive practice she is engaged in transparent. This awareness, which determines action-planning (erasing the text or keep it), is emotionally motivated. The tension she felt when the message was shown in the past emotionally tuned her in a certain way in the current situation, and made the window showing the message particularly interesting at that moment. She was prepared to deal with the emotional consequences of her actions, and then to act in a certain way according to those emotions and concerns. In this sense, it can be said that relevant affordances in a given context are *affectively enacted*: salience is brought forth by the agent's reactions caused by an affective tone.

This explanation of selective responsiveness is particularly interesting for the purposes of this chapter. Indeed, contrarily to Clark's explanation, it seems to have many pros.

i) It is phenomenologically viable. Indeed, it accounts for what agents experience in their everyday practices, and it does so by taking into account many variables, such as the cognitive agent's embodied condition, emotions, actions, concerns, and so on.
ii) It is epistemically economic, because it relies on the twofold nature of intentionality only. This one is conceived as practical and affective.
iii) It does not make use of unclear and unwarranted concepts (i.e. AORs).
iv) Moreover, it fits with experimental results from cognitive neuroscience concerning emotion regulation and impulse control

> (Frijda, Ridderinkhof, Rietveld 2014). Indeed, it has been shown that dopamine[92] neurons are sensitive to changes in the *prediction* of the "goodness" of ongoing actions. Contrary to what was the mainstream account of the role of dopamine in action-control - according to which dopamine reinforcement-learning is said to mediate feelings of satisfaction experienced by an animal receiving a reward (and then dopamine was considered as an "outcome" of the emotional appraisal of an action already performed)- recent research (Wickelgren 1997) showed that once the animal has learned to perform the task correctly, the presentation of rewards does not influence the dopamine system. On the contrary, the dopamine system becomes active in the anticipation of a forthcoming reward rather than on the delivery of the reward itself (Holroyd, Coles 2002: 681-682).
>
> This seems to provide evidence for the hypothesis -formulated in Merleau-Ponty 1945, Varela 1999 and endorsed in Rietveld 2008a- that emotions work as control parameters for action, and then contribute to the goal-directedness of the agent's practices, in which the motivation of selective responsiveness to affordances was individuated.

Nevertheless, even if this account of selective responsiveness has many pros, and then I claim that is a better explanation than the representational proposal, it seems to me that, in order to be more complete, some integrations are required. Indeed, this explanation is particularly focused on *one* correlate of the affordance-relation. It explains how some affordances become relevant in a given context, and then how they are adequately grasped by the agent, by stressing on the agent's characteristics (her motor and emotional intentionality).

Obviously, this does not mean that the role of world in this dynamics is neglected. As explained previously, Merleau-Ponty's notion of intentionality this non-representational account of selective responsiveness relies on considers the world as "the immanent term" of the agent's practical interactions. This means that the world in practices is not an object, an ideal

[92] Dopamine is a brain neurotransmitter, identified about fifty years ago, and it is correlated to emotions. Indeed, it has been shown that there is a direct evidence of dopamine involvement in emotional recognition processes: emotional processes are dependent upon many brain structures, the majority of which form part of the limbic system, which is subjected to dopamine innervation (Salgado-Pineda, Delaveau, Blin, Nieoullon 2005: 228).

term that gives rise to an indefinite variety of perspectival views, remaining equivalent although the perspective changes, but it is rather a flexible component of the intentional relation. Think again about the football player's example. Each maneuver performed by the player "modifies the character of the field and establishes in it new lines of force in which the action in turn unfolds and is accomplished, again altering the phenomenal field" (Merleau-Ponty 1942/1963: 168). This suggests that the world is a constitutive part of a situated intentional movement. The explanation does not neglect the world, but it rather thinks about the way it melts into the cognitive agent's activities.

Nonetheless, it seems to me that this explanation does not develop an analysis of the changes that the agent's selective responsiveness produces in the affordance landscape itself, nor it focuses on the very role of the world in selective, contextual responsiveness. Rietveld and Kiverstein consider this point very quickly when they talk about "education of attention", mentioned by Ingold, and they just say that the acquisition of new skills leaves "landmarks" that guide the agent's activities (Rietveld, Kiverstein 2014: 331). Or again, they write: "it is the world, the soliciting relevant affordance encountered in a concrete situation, which motivated an individual agent to do one thing rather than the other" (Rietveld, Kiverstein 2014: 345), adding that (higher) cognitive capacities should be thought "in terms of skillful activities in sociocultural practices and the material resources exploited in those practices" (Rietveld, Kiverstein 2014: 326). My point is that, despite they acknowledge the importance of the socio-material world in their account of selective responsiveness, they do not explore this idea further.

In order to develop Rietveld and Kiverstein's insights about this point further, I will try to flesh out the affordance-based explanation of contextual relevance by integrating it with Peirce's concept of indexicality[93]. To develop my explanatory hypothesis, I make the premises of my argumentation clear.

Recent approaches to cognition (Hoffmeyer 2015; Maran, Kull 2014; Morgagni 2012; Nöth 2011; Nöth, Kull 2001; Sonesson 2012; Zlatev 2012; Windsor 2004), which integrate semiotics with other disciplines - such as theoretical and applied biology, landscape studies, communication studies, cognitive sciences, phenomenology, and the humanities more in general-, explain cognitive phenomena as unfolding in the relation between organisms and environment. This relation is said to be mediated by signs.

According to this approach, not only cognition is a meaning-making process (as Peirce already pointed out) embodied in signs with different features, but the environment in which meaning-making processes develops is a *semiotic*

[93] To understand better the general theoretical link between ecological psychology and Peirce's semiotics at the level of secondness, where indexes lay, see O'Neill 2008: 155.

niche (Atã, Queiroz 2016). Semiotic niches are conceived as cognitive niches, namely as cognitive spaces or habitats which are the product of human and non-human animals' natural and cultural evolution (Laland, Olding-Smee, Feldman 2000). Those habitats are particularly helpful for the development of agents' cognitive practices. Indeed, as explained in my discussion of the concept of Umwelt, they work as external cognitive scaffoldings, offering saliences that can be externally manipulated by agents. This makes the agents' internal cognitive load lighter.

A very simple example of cognitive niche is the following one. Think about "a novice bartender [who] inherits an array of differently shaped glassware and cocktail furniture, and a practice of serving different drinks in different kinds of glass" (Clark 2005c: 256, text into brackets added). The novice bartender inherits a structured environmental space from the expert bartender who used to work at the coffee shop. Indeed, the expert bartender lined up differently shaped glasses in a spatial sequence corresponding to the temporal sequence of drinks orders. The problem of remembering what drink to prepare next is made easier by this organization of the space, because it is translated into a simple perceptual problem. In this sense, it can be said that the spatial organization of the objects and tools the bartender uses can be seen as a cognitive niche. It is something that was created in time, thanks to the practice of many bartenders, and it is inherited by the new barman, which learns how to be effective in her practice relying on a pre-structured niche, whose efficacy was tested in the past.

Now, the mentioned semiotic approaches to cognition holds that cognitive niches are populated by different kinds of signs. That is to say, animals structure their environment in a certain way. In doing so, they also learn to see them as populated by different kinds of signs (embedded in material artifacts, gestures, intersubjective dynamics), which afford a great variety of actions and practices (Malafouris 2013: 89-90; 94) in virtue of their semiotic properties. To put in other words, in their cognitive practices, agents perceive a landscape of signs that functions as a "semiotic interface between resources and organisms" (Farina 2008:76; see also Farina, Belgrano 2005). This one allows them to exploit the environmental resources in a successful way.

For what concerns the problem I am dealing with in this chapter, namely contextual relevance, I think that there is a kind of sign that can be taken into account in order to understand why some signs stand out more than others in the rich landscape of signs: the indexical sign. My claim is the following one: indexical signs, in virtue of their semiotic constitution, are signs that stand out in the landscape of signs, they trigger the agent's attention, and make her acting in a certain way. This is to say that the semiotic constitution of these signs

explains *how* and *why* agents are attracted by some features of the unruly and noisy environment and not by other features.

To explain my point better, I consider Peirce's definition of index.

The term "index" shows up in Peirce's classification of signs in relation to the dynamical object[94] they refer to (CP 8.335). In the *Baldwin Dictionary* (1901), Peirce writes:

> "[Index:] A sign, or representation, which refers to its object not so much because of any similarity or analogy with it, nor because it is associated with general characters which that object happens to possess, as because it is in dynamical (including spatial) connection both with the individual object, on the one hand, and with the senses or memory of the person for whom it serves as a sign, on the other hand. [...] Indices may be distinguished from other signs, or representations, by three characteristic marks: first, that they have no significant resemblance to their objects; second, that they refer to individuals [...]; third, that they direct the attention to their objects by blind compulsion. [...] Psychologically, the action of indices depends upon association by contiguity, and not upon association by resemblance or upon intellectual operations". (CP 2.305 – 2.306)[95]

An indexical sign[96] is a sign that represents an individual object i) by entertaining a dynamical connection with it, and ii) by entertaining a dynamical connection with the agent's perceptions or memories of that object, namely it entertains a deictic connection with *bodily* states[97] (Stutz 2014: 5).

This connection between the object and the agent's perceptions and memories -which, as already said, is a "direct physical connection" (W5: 254)- is a *dynamical* connection. In fact, Peirce says that the effects indexes produce are compulsory *reactions* of attention. These reactions i) change the agent's way of being situated in her practices, by directing her attention

[94] Peirce distinguishes between the Immediate Object, "which is the Object as the Sign itself represents it [...], and the Dynamical Object, which is the Reality which by some means contrives to determine the Sign to its Representation" (CP CP 4.536). The Immediate Object is the object as it is represented in the semiotic relation, the dynamic object, is an object of world.

[95] For a different discussion of this Peircean citation, see Fumagalli 1995: 338-345.

[96] Here I use the expression "indexical sign" instead of "index" as an opportunity to make an aspect of Peirce's explanation of indexicality more clear. I talk about "indexical signs" because, according to Peirce, there are not pure indexes, but signs whose indexical features are very strong (CP 2.306). Therefore, when I will use the word "index", remember that I refer to signs whose indexical features are more predominant than the iconic and symbolic ones.

[97] Peirce says that Indexes have a "real *physiological force* over the attention" (CP 8.41; emphasis on the word "physiological" added).

towards a specific aspect of the practice, and, in doing so ii) they make her ready to do something upon that new perceived salience.

This is to say, an index is a sign that represents its object as an *individual salience*, and it does so by catching the agent's attention brutally, by pushing her to do something. What the agent who reacts to indexes does is exactly that kind of action the indexical mark requires. To the extent that indexes are in a direct connection (or contiguity) with the object they represent, their cognitive effects are tied to the particular context in which they show up: they "light up" an object of the semiotic landscape as the individual salience the attention has to be directed towards compulsorily.

To explain this point better, I consider some examples of index given by Peirce.

i) The term "here" in the sentence "Within a thousand yards of here" (CP 2.305). "Here" draws the speakers' attention towards the ground between them. It is a sign that makes sense in that given context, in the presence of that portion of ground. The effect it produces is "pointing energetically to the ground", making the agents feeling that portion of ground that was not salient in their practice before "here" was uttered. Crucially, it does so "brutally", namely without entailing a "description" of the object indicated[98]. It makes the object directly present to the agent's attention.

ii) A finger that points to the fire (CP 2.305). The finger functions as an index because it is dynamically connected to the fire. It is like a fire alarm that forces the agent's eyes to turn that way, by producing a rupture in the "normal" flow of signs the cognitive agent's experience unfolds by. The pointing finger entails a compulsion to see the *hic et nunc* (CP 2.336) of the practice the agent is engaged in. At the same time, it marks an individual element of that practice.

iii) A "weathercock", which takes the same direction of the wind (CP 2.286), and forces the agent to look in the same direction it points to.

Now, these examples seem to suggest that there are some "things" in the environment (a word uttered during a linguistic practice, a gesture, an object) that solicit the agent's attention more than other things do. As already

[98] See CP 3.361, where Peirce says that Indexes assert nothing.

explained, according to a semiotic approach to cognition modeled on Peirce's thought, cognitive practices *in general* can be said to develop through signs, which afford some kind of action. The notion of index seems to suggest that agents are prone to act upon certain given signs and not upon others because there are signs, namely indexical signs, whose *demand character* is stronger. When I say that their demand character is stronger, I do not want to mean that *only* indexical signs push agents to react or act in the semiotic landscape, or that indexical signs are the only ones able to show the practical bearings of an externally represented object. Rather, I want to suggest that indexes, in virtue of their high degree of context-sensitiveness (that is their being physically connected to the *individual* object represented, and physiologically connected to attentional phenomena), and in virtue of their ability to solicit immediate reactions, function as *markers of salience*. They stand out in the landscape of signs by forcing the agent's attention towards what should be done.

That is why I claim that indexes explain how and why some affordances matter more than others, integrating the enactive approach to "contextual relevance".

On the one hand, because indexes stand out in the landscape of signs, they put other signs in the background for a moment. This would explain how some features of the world, in virtue of their constitution (in this case, semiotic constitution) matter more than others in a given situated practice.

On the other hand, since they are dynamically connected to the object represented in the semiotic relation and dynamically connected to the agent's attention, they are good epistemic candidates to explain how and why agents usually perform the right kind of action in perceptually undetermined environments. Indeed, in indexical semiotic relations, the agent's attention is brutally triggered, and the agent is forced to do exactly what the indexical dynamics ask to do. Indexes ask to react to that individual object (the environmental salience) they represent.

To put in other words, indexes seem to do the job of action-oriented representations, without entailing the problems these epistemic posits have in Clark's explanation.

First, they are tailored to the agent's attentional experience. They are *directly* connected to the agent's attentional phenomena because they operate in the realm of physiology: they trigger and immediate reaction.

Second, they account for the action-command aspect ascribed to AORs, since the force they have on the agent is brutal: they make her acting in a certain way compulsorily.

Third, they are tied to the present context of the agent's practice. In particular, they explain how that particular individual environmental feature marks a salience that will determine how the cognitive practice will develop.

Crucially, all the job indexes do is performed outside of the skull. Instead of being embedded in internal representational mechanisms, indexes are part of a semiotic relation that directly takes place in the flow of the cognitive practice. In this sense, indexes not only account for the antecedent of action, working as triggers for reactions, but also account for action *in* indexical practices: the agent's reaction they trigger is part of the semiotic relation itself.

That is why I said that Peirce's concept of indexical sign avoids Clark's AORs problems. Indexes are not internal signs primarily, then they do not entail an internalistic prejudice. Moreover, they seriously account for action in semiotic practices, not conceiving of it as an output of internal representations.

Furthermore, this externalist, active, and dynamical explanation of salience in context seems to be consistent with the general ideas suggested by the enactive explanation. Indeed, it does not presuppose the complex internal representational mechanism rejected by the enactive approach to cognition, and it does not entail conceptualization (Short 1996: 489): indexes are part of the agent's pre-reflective engagement with the world. Moreover, indexicality explains the agent's meaningful "being-in-the-context" as immediate -as the enactive explanation does- because indexes are said to trigger the agent's attention directly and immediately.

Given that the two explanations –at a broad theoretical level, at least[99]- seem to be consistent the one with the other, I claim that Peirce's indexes are good candidates to make the enactive approach to selective responsiveness more complete. As already said, the enactive explanation acknowledges the importance of the role played by features of the world in explaining how and why some affordances matter more than others, but it does not develop this point further[100]. Conceiving of relevant affordances as indexically grounded would fill this gap. Indeed, it would explain selective responsiveness as determined by i) the agent's affective and practical intentionality, and by ii) some aspects and dynamics of the environment which, thanks to their indexical nature, stand out in the rich landscape of affordances, controlling action immediately and compulsorily (West 2016: 225) "from the outside".

[99] See Gallagher 2009.

[100] For a more general semiotic critique of the Enactivist approach to cognition concerning the primacy given to human agents in cognitive practices see De Jesus 2014.

II.7 Concluding remarks.

This chapter was devoted to discuss the problem of practical knowledge by considering the issue of selective responsiveness to stimuli (i.e. the Generalized Frame Problem) as a special case of this problem. I showed that both the EMH and the Enactive approach to cognition pay particular attention to practical knowledge, and they explain agents' smooth engagement with their world by focusing on the concept of activity. Nevertheless, these two approaches to cognition disagree on the role played by internal representations in these practical engagements.

The EMH wants to preserve internal representations, and it reassesses them in order to make them matching an embodied, active, and extended approach to cognition. On the contrary, the Enactivist explanation wants to get rid of internal representational mechanisms endorsed by EMH.

In my discussion, I explained why I think that Clark's account of selective responsiveness is not philosophically viable, and I did this by considering some objections to AORs. These ones showed that AORs do not fulfill some of the criteria necessary to talk about representations properly. Therefore I claimed that the use of a representational gloss in the case taken into account is unwarranted.

Moreover, I pointed out that the way Clark talks about AORs in his book *Being there* offers other reasons to think that the EMH is not as radical as one could think. The way this explanation deals with AORs show that the EMH is committed to an internalistic prejudice. Indeed, instead of describing the agent's practical engagements with the world by stressing on practice itself, Clark's explanation seems to consider practice as a derivative mode of cognition forerun by internal mechanisms, which seem to be considered to be the original core of cognitive processes. Furthermore, even if those mechanisms are said to be endowed with active features, the way they entail action is too weak. Indeed AORs have a conative power, but they are not actions in themselves.

Considered the problems Clark's explanation entails, I tried to find a possible solution to the Generalized Frame Problem by looking at the way the enactive approach to cognition deals with selective responsiveness and contextual relevance. I claimed that this explanation is more viable than Clark's one because it accounts for phenomenological experience, it is epistemically economic, it does not make use of an unwarranted epistemic posit (the concept of AOR), and it fits with experimental results from cognitive neurosciences.

Despite I am sympathetic with the enactive explanation, I also claimed that this approach to contextual relevance should be integrated in a broader explanatory framework, in order to make it stronger and more complete.

Indeed, despite explanations that cross ecological psychology with Merleau-Ponty's phenomenology acknowledge the importance of the socio-material world in episodes of selective responsiveness to affordances, they not develop this point further. This because they do not take into account in detail how certain structures of the environment display salient features in virtue of their constitution.

Therefore I suggested to integrate the enactive explanation within a semiotic explanatory framework inspired by Peirce's thought. In particular, I suggested that cognitive niches can be seen as semiotic niches, and that saliences in semiotic niches can be interpreted as indexes.

Conceiving of relevant affordances as indexes is helpful for two reasons.

Firs, at a general level, Peirce's semiotic account of cognition is suitable to explain cognitive phenomena from an externalist perspective, and it does so by focusing on the concept of activity interpreted in a semiotic way. This is precisely what the non-representational explanation of selective responsiveness looked for.

Second, by focusing of semiotic features of cognitive niches, the theory can explain why some affordances matter and stand out more than others, paying attention not only to the agent's abilities and concerns, but also to the other parts of the coupled system made of the organism and its cognitive niche.

Therefore, I suggest that this integrated account of the Generalized Frame Problem is able to provide a balanced, non-internalist, non-agent-bounded, and active explanation of adequate cognitive performances in context.

Moreover, this semiotic implementation of the enactive approach to cognition seems to be suitable to account for the mind as an *extensive process*, claim endorsed in Chapter I, and that, contrarily to the EMH, is not compatible with internalism at all (Hutto, Kirchhoff, Myin 2014: 4). According to this explanation, the mind should not be thought to be an internal capacity that just exploits external scaffolding, or that sometimes is instantiated by external resources. Extensive minds are *always* and *already* world-involving (Hutto, Myin 2013: 137).

Peircean semiotics, as showed in my discussion of selective contextual responsiveness, fosters this idea. This for two main reasons:

i) As "extensive enactivism" does (Hutto, Kirchhoff, Myin 2014; Hutto, Myin 2013), it gives an active account of cognitive processes, and it does so thanks to the concept of sign-action (semiosis).

ii) It offers interesting conceptual tools (e.g. indexicality) to examine the way the world is involved in cognitive practices. In this way it helps the enactive proposal to flesh out its conception of the mind as *already* world-involving, since it does not conceive of salient features

of the world as dependent on the agent's concerns only, but rather as signs that stand out in virtue of their semiotic constitution.

Chapter III

Predictive Processing.
The problem of representations

III.1 Introduction. Approaching Predictive Coding.

In this chapter, I will consider Predictive Coding, a contemporary explanation of action-perception loops, whose explanatory unit –in particular in its classical version- consists in the brain's anticipatory mechanisms. In my exam of Predictive Coding, I will examine the way the EMH and the enactive approach to cognition interpret the core ideas of this theory.

To make my core claim clear, in §III.2, I contextualize Predictive Coding in a broader theoretical frame, explaining why I think that this approach to action-perception loops is aimed at discussing the same issues considered in Chapter II: anticipatory and recognition processes postulated to explain agents' smooth engagement with their environments. Therefore I formulate a working hypothesis, which will guide my argumentation: if a version Predictive Coding relies on the same concepts considered in the previous chapter, in particular AORs, the explanation will probably entail the same philosophical problems considered in my previous discussion about practical knowledge.

After this brief theoretical contextualization of Predictive Coding, in §III.3, I will take into account a strongly internalist version of this approach to anticipatory mechanisms: Hohwy's "Predictive Mind". The exam of Hohwy's theory has two main purposes: i) introducing the key concepts of Predictive Coding, and ii) preparing the discussion of Clark's approach to this theoretical framework.

In §III.4, I will examine Clark's approach to the "predictive mind", and I will discuss the way he tries to make this theoretical framework matching the philosophical requirements of the EMH. I will claim that Clark's strategy to extend the "predictive mind" is too weak, because it makes use of concepts that entail internalistic prejudices (Parity Principle and sampling).

In §III.5 I will develop my objections further, by taking into account the way Clark makes use of the notion of representation in his "Predictive Processing". My exam of the concept of representation will show that the use of notions such as "probabilistic representation", "error representation", and "action-oriented representation" is philosophically unwarranted, because the processes Clarks wants to explain do not meet the conditions that philosophers of cognitive sciences endorse to talk about representation properly. Moreover, I will discuss the concept of "action-oriented representation" in more detail, and I will claim that the appeal to this notion is a symptom of how Clark underplays the role of embodied action in predictive mechanisms. Explaining real embodied, contextual action through the notions of "action-oriented model" and AOR, Clark fails to meet the requirements of a more enactive account of

cognition. This makes the "peace treaty" he suggests enactivists to sign by endorsing "Action-Oriented Predictive Processing" unacceptable from an enactive perspective.

Given the problems Clark's proposal entails, in §III.6, I will take into account another way to conceive of predictive processes, developed within an enactive framework that couples the already discussed work on the "field of relevant affordances" and some insights form "enactive affective science" (e.g. the notions of situated core relational theme and enactive appraisal).

In §III.7, I will sum-up my considerations about the way the EMH and the enactive approach to cognition deal with the issue of prediction in cognitive processes. Moreover, I will notice that the centrality of the affective dimension in the enactive proposal not only helps the explanation to give a more situated account of PP, but it is also crucial to prepare the philosophical ground to answer an objection against the enactive approach to cognition: the idea that enactivism is a new form of behaviorism because it rejects internal representations and computations.

III.2 Setting out the problem: the theoretical continuity between AORs and Predictive Processing.

In the previous chapters, I took into account the broad project that "4Es" approaches to cognitive sciences promote. I pointed out that, looking at the broad landscape of those theories, theoretical proposals that fall under the label "4Es cognition" or under the umbrella term "embodied cognition" (Violi 2003: 59-62) point to "mind extension", or at least they aim to reconsider the way we should do cognitive sciences, by reconfiguring or rejecting cognitivism (Menary 2010d: 459). Indeed, the overall project of those approaches to cognition seems to consist in a redefinition of the mind's boundaries through the inclusion of the role of the anatomical, sensorimotor, lived, and affective body, and of the natural, cultural, and social environment in cognitive processes.

Nevertheless, I also noticed that the inclusion of those factors - which I previously called "dirty factors" to mark their contrast with cognitivists' sanitized explanations of cognition (see §I.2) - does not guarantee the development of the *full thesis* of "mind extension". As I explained in chapter I -and as other scholars have already claimed (e.g. Fusaroli, Paolucci 2011; Gallagher 2011a), the progressive inclusion of those factors in contemporary explanations of cognition and the mind gives rise to different pictures of the mind, whose departure from cognitivism is more or less radical.

In Chapter I it has been shown that, despite works in cognitive linguistics, "grounded cognition", and some versions of "embedded cognition" make great steps forward in the project of understanding the mind as embodied and enhanced by external scaffoldings, those approaches do not seem to explicitly draw a *systematic* picture of cognitive processes aimed at rethinking the very *ontology of the mind.*

On the contrary, this theoretical project is at the core of two specific versions of embodied cognition: the enactive one and the functionalist approach to embodiment (Gallagher 2011a; Gallagher 2017: 35) endorsed by Clark in his EMH. In my discussion of the two theories I noticed that, despite these two approaches share the idea of "mind-extension", there is a huge theoretical difference between them. This not only because they come from different philosophical lineages (phenomenology, theoretical biology, ecological psychology vs. analytic philosophy and AI)[101], but also and especially because

[101] Here it is worth to notice that this individuation of the sources of the enactivist approach to cognition and the EMH works at a broad level (Gallagher 2017: 48), and it is accurate if we consider early works on enactivism, which clearly emphasized their affiliation with the phenomenological tradition, in particular with Merleau-Ponty's work. Indeed, some of the more recent enactive approaches to cognition (e.g. Hutto, Myin 2013; Hutto, Myin 2017) are

their departure from classical views of the mind (i.e. representational approaches, or internalistic views of the mind) is *more or less radical*.

Even if both approaches -to different extents- seem to introduce themselves as "radical theories" in the contemporary debate (see for example Chemero 2009; Clark 2008: 68, 105; Clark 2015c; Hutto, Myin 2013; Hutto 2015; Gallagher 2011a and Gallagher 2017: 40 for the label "radical embodiment"; Varela, Thompson, Rosch 1991: xvii, 150), a more careful analysis of these philosophical stances shows that the revolutionary theoretical shifts those theories talk about are sometimes just prospected, and not actually developed in the practice of their philosophizing. To put in other words: even if both the enactive approach to the mind and the EMH, *at first glance,* seem to radically differ from classical cognitivism and internalism, a comparison between these two views of cognition shows that one approach is more suitable to develop a theory of "mind extension"[102] than the other one.

Surprisingly, the Extended Mind Hypothesis –whose label sounds like is meant to be a clear statement of the full thesis of "mind-extension"- still sticks to a conservative view of the mind, I claimed.

This point was briefly taken into account in §I.9. There I expressed my doubts about the label "extended mind", and about the theoretical value of the parity principle, explaining why I think that these two points hide an internalistic prejudice. Furthermore, I pointed out how the centrality of the concept of mental state -at least at a general level- makes the theory unable to account for the *active process* (Gallagher 2011b; Gallagher 2013) through which the mind unfolds in the inter-action between agent and environment.

These considerations were developed further and in more detail in Chapter II. There I focused my attention on a specific problem: the way the EMH and the enactive approach to cognition account for the agents' smooth engagement with the environment, selectively responding to environmental saliences, and then displaying a behavior adequate to the context of the cognitive practice.

The overall structure of the argumentation was the following one. If it is true that at a general level the EMH and the enactive approach to cognition seem to point to an "extensive picture" of the mind, if one considers the details of their individual arguments about specific problems (namely, if one analyzes each theory at work), it can be shown that the EMH still gives a conservative explanation.

---

not committed with the phenomenological approach and with the life-mind continuity thesis (Ward, Silverman, Villalobos 2017: 372), and adopt a more analytical philosophical style.

[102] The idea that the mind is an interactive process made of the transactions between agents and their environments.

Considering the particular case taken into account, I claimed that the appeal to AORs to explain selective responsiveness to environmental saliences -issue that is part of the broader topic of practical knowledge, conceived as the ì looping of action and perception (see §II.2)- is a special case of "internalism at work". In fact, even if Clark's broader philosophical project aims at considering seriously how cognitive agents are successful at performing cognitive practices looking at the way external props shape the cognitive practice, when Clark considers action-perception loops claims that those processes are controlled by *internal cognitive resources*: action-oriented representations.

Those ones, according to Clark, should represent a step forward in our philosophical explanations because -instead of being abstract representations that have nothing to do with the body in action and with the individuality of the environmental feature represented- account for the agent's body and for the specificity of the action that should be performed. Indeed, they are said to represent a specific motoric command that anticipates what should be done in the context of the cognitive task.

I suggested that this way of considering action-perception-based cognitive practices is an example of a hidden attempt to conceal a conservative attitude towards the mind by means of philosophical statements (i.e. AORs radically differ from cognitivists' representations; Clark 1997a: 49) that are just apparently progressive. If one the one hand it is true that Clark actually tries to give a body to internal representations, and to put action in representation, on the other hand it is also true that the way this philosophical proposal is developed is not successful. In fact, a semiotic analysis of Clark's AORs shows that, in his account, action keeps on laying outside representational mechanisms. Indeed AORs prepare or anticipate action. This is to say that, in Clark's explanation, action is not a component of representations in themselves, it does not actually define the very concept of AOR. On the contrary, it is an element considered in the broader economy of the explanation and *then* extrinsically attached to the label "representation". Moreover, the representational mechanism itself is conceived as static in its individual components (see the objections considered in §II.5). The explanation deals with theoretical entities (AORs) that are similar to unconscious dispositional states, namely AORs are identified with internal *states* that have just conative features, and not real active features.

That is why I looked at the enactive account of selective responsiveness and contextual relevance (Bruineberg, Rietveld 2014; Rietveld 2008a; 2008b; 2012a), hoping to find there a better account of the philosophical issue considered, namely an explanation that does not entail the problems Clark's

approach suffers form. In those enactive works, I actually found a more externalist, active, dynamical and not intellectualistic account of the specific problem taken into account, which meets the requirements of a genuine theory of "mind-extension".

First, considering this proposal at a broad methodological level, it can be said that it extends the explanation by introducing factors that Clark does not consider (e.g. affects that anticipate or structure action-perception loops).

Second, the enactivist proposal gets rid of internalistic prejudices, by cleaning up its explanation of action-perception loops from internal action-oriented representations. This idea -I suggest- is a fundamental step for the project of mind-extension. Indeed, by substituting AORs with "relevant affordances", the explanation actually gives space to the environment[103] in its consideration of "meaningful behaviors". To put it roughly, the core idea of this approach- already expressed by James Gibson (Gibson 1979/2015: 215)- is that behavior is regular even it is not regulated by complex internal structures. Behavior is regular because agents perform their actions in a rich field of regularities that triggers agent's actions.

Furthermore, contemporary enactive approaches to cognition explain why agents usually deploy the correct kind of action in a given context. Thanks to their practical and affective intentionality, agents structure their environments as disseminated of "attractive saliences". This allows them to cope with the environment smoothly because the affective tone the environment shows up by makes some invitations for action standing out more than others. In this way, the enactive approach not only accounts for the perceptual mechanisms that trigger agents' actions, but also accounts for the preparatory stages of this cognitive activity. Crucially, it does so without making use of concepts such as internal model and internal representation (concepts that, for what concerns action and perception, has been shown to be philosophically unwarranted). On the contrary, it conceives of the non-intellectualistic dynamics that structure both the agent and the environment (making them attuned the one with the other) looking at the affective dimension.

[103] Notice that if it is true that the enactive explanation of contextual responsiveness gives more space to the environment than Clark's one does, it is also true that it does not give a detailed account of features of the environment this phenomenon of selective responsiveness depends on. That is why, in II. §6, I suggested that a semiotic integration of this explanation grounded on the idea that relevant affordances have indexical features can help make this account stronger.

What I will do in this chapter goes in tandem with what I did in the previous one. As a matter of fact, the structure of my discussion will be nearly the same. I will consider another specific point these two approaches to cognition disagree on, in order to understand if the ways these theories deal with that specific problem promote or do not promote a genuinely "extensive" account of the mind. The problem I will consider in my discussion is the way the enactive approach to cognition and the EMH deal with a contemporary model of the brain: Predictive Coding.

I chose to talk about this new approach to brain functions for at least three reasons.

First, both supporters of the EMH and enactivism are showing a huge interest in this new approach. According to contemporary approaches to cognitive sciences, Predictive Coding is an appealing theory that can be discussed and integrated in their philosophical projects. Indeed, not only it provides a detailed account of how the brain works, but it also offers philosophical suggestions useful to think of "a meeting point for the best of many previous approaches" to cognition (Clark 2016: 10) or of a "mature science of the embodied mind" (Clark 2015b: 16). Indeed, a theoretical discussion of this neuroscientific model of explanation can be useful to think about open questions in philosophy and theoretical cognitive science.

At a general level, looking at how the brain works can offer chances to rethink about "the form and structure of the human experience itself" (Clark 2016: 3), philosophical issue at the core of second wave-cognitive sciences' project (see §I.2).

Moreover, Predictive Coding deals with more specific issues at the core of the contemporary debate in cognitive sciences. Above all, it offers the opportunity to question the status of perceptual mechanisms, putting on the table an old issue about the nature of perception: is perception inferential or direct? This problem, as implicitly suggested in Chapter II, was at the core of Gibson's project of ecological perception. Ecological accounts of perception, according to which perception is direct, inspire many contemporary approaches to cognitive sciences, in particular the enactive one. Predictive coding, widely relying on the concept of inference in its explanation of perception, seems to offer neuroscientific objections against this direct approach to perception. Given this difference concerning the core concepts of the two theories, is it possible to think of an integration of this scientifically accepted model of perception and philosophical explanations such as enactivism (Clark 2012; Orlandi 2016)?

This philosophical question about perception and inference, I suggest, asks us to look back at the problem of action-oriented representations: this is the third and main reason why I think it is worth to consider Predictive Coding right after an exam of the problem of AORs. My hypothesis is that there is a strong philosophical connection between philosophical discussions of Predictive Coding (i.e. perception as inference) and the debate between representational and non-representational accounts of action-perception loops, in which the concept of AOR is pivotal.

I think that there are two reasons that motivate this connection.

First, as explained in Chapter II, the debate about AORs concerned different ways of explaining action and perception. Minimal representationalists claim that action and perception are mediated by internal action-oriented representations. This is to say, the core idea of (minimal) forms of representationalism is that perception and action are mediated by internal representational entities. Predictive Coding, with its inferential account of perception, seems to point to the same idea of "mediated perception" (Gładziejewski 2016), while enactive and strongly affordance-based accounts of perception suggest a radically different story.

Second, the core ideas of Predictive Coding -and, as I will explain later, of philosophical reassessments of this theory- seem to be already present in the debate about AORs.

Basically, according to this approach rooted in neurosciences, our mind's ability to perceive the world -that is the ability of making sense of the manifold sensory inputs coming from the senses- can be explained looking at predictive mechanisms of the brain (Hohwy 2013: 1). The core idea of this theory is that agents *smoothly* and *quickly* (Clark 2015b: 21) engage with their world by means of perception (that is continuous with action and attentional mechanisms; Hohwy 2012; 2013; 2014) because evolution and neurodevelopment endow brains with hierarchical statistical models of the causes of sensory inputs (Howhy 2014: 4; 63). Those statistical models allow the brain to successfully predict or anticipate its sensory states. Instead of just passively receiving sensory inputs, the brain builds up perceptual data by matching what it already "knows" about the world and the cognitive agent's body (Clark 2013a) with the novelty coming from the senses. This suggests that perception and action are shaped by expectations derived by previous knowledge and that the more this set of cognitive priors is accurate (i.e. the more prediction errors are minimized), the more cognitive performances will be successful. The more the brain's anticipations of its own states are successful,

the more the agent's behavior will be attuned to the environment where the cognitive practice takes place.

This description of our brains as "pro-active prediction systems" (Clark 2015b), I claim, is not that far from some explanations of action-perception loops taken into account in Chapter II. Indeed, the core idea of this approach to perception and action seems to be this one: agents are able to engage successfully with their worlds because action-perception loops are guided by internal models that have an *anticipatory* power. Action-oriented representations –at least in Clark's version given in *Being there* (Clark 1997a)- seem to have the same function: they are said to guide behavior by representing features of the environment as contents of an action-command. Then, what is represented is i) an anticipation, or a simulation (Clark, Grush 1999) of what the agent is supposed to do in the environment in the proximal future, and ii) an anticipation of the sensorimotor states that will occur during action[104].

This theoretical similarity between Predictive accounts of our brains and the problem of AORs is particularly striking in Clark's latest book: *Surfing Uncertainty. Prediction, Action, and the Embodied Mind* (Clark 2016). Clark's book aims to integrate the Predictive Coding model in a broader perspective, consistent with his EMH. To do that, he basically works on Predictive Coding models to make them matching the requirements of an *extended* and more *active* view of the mind. Clark's argument is twofold.

First, he states that "nothing in the PP [Predictive Processing] framework materially alters [...] the arguments regarding the possibility of genuinely extended cognitive systems" (Clark 2016: 260, text into brackets added). This because the agent's body and the world -namely biological and non-biological external resources- provide additional strategies to minimize the brain's prediction error mechanisms, which are at the core of good anticipation strategies. Long-term material and socio-cultural environmental structures, and the actions we perform in this structured world, minimize prediction-error and improve the anticipatory power of cognitive systems. By structuring the world in a certain way -that is in a way suitable for the distribution of the cognitive load among different resources- we make our sensory predictions coming true

[104] Notice that Clark himself draws a comparison between AORs and Predictive Coding models in his article "Whatever the next? Predictive brains, situated agents and the future of cognitive science", when he says that Grush's "emulator theory of representation" (Grush 2004) shares many features with the predictive processing story (Clark 2013a: 18; 24, footnote 44).

(Clark 2013a: 21), allowing the brain "to do as little as possible while [...] still solving the problem" (Clark 2015c: 12).

Second, he takes into account the role of action in prediction by talking about something that recalls his previous work on AORs. In fact, Clark defines his way of conceiving of PP as "Action-Oriented Predictive Processing" (Clark 2013a; Clark 2016). At the core of this view of Predictive Processing there is the idea that the job the brain does (i.e. predicting its own states and reducing mismatches between top-down predictions and bottom-up flows of information) is about "selection and control of world-engaging action" (Clark 2016: 250). This talk about *selection* and *control* of world-engaging action –I claim- clearly echoes what has been considered in Chapter II, namely the Frame Problem, to which action-oriented representations were supposed to find a solution. Indeed, as previously explained, AORs are conceptual entities that are supposed to help philosophers to explain

a) why agents display that right kind of action in a given context (selective function), and

b) they work as internal controllers for action (control function).

This is to say, Action-Oriented Predictive Processing and Clark's Action-Oriented Representations, although they are presented as distinguished theoretical entities, seem to be introduced in Clark's explanation to account for the same problem: explaining how behavior is controlled, and how and why agents display the correct kind of behavior in a given context.

Furthermore, the way Clark talks about this function of selection and control performed by Action-Oriented Predictive Processing seems to be consistent with the general terms in which the debate about AORs is developed. Clark describes this problem talking about "affordance competition" (Clark 2016: 251; see also Cisek 2007 for Clark's primary reference). The idea is that agents' brain continuously computes multiple probabilistic inflected possibilities for action to make conscious experience successful in the environment. It seems to me that this is a new way (where the novelty is the introduction of the probabilistic nature of internal mechanisms) to talk about the problem of relevant affordances taken into account in Chapter II, where I discussed the enactive approach to selective openness to the world[105].

[105] A connection between these two proposals is explicitly suggested by the authors of the enactivist proposal considered in Chapter II (Kiverstein and Rietveld 2012), in the article "Dealing with context through action-oriented predictive coding. A commentary in Clark's 'Whatever the next? Predictive brains, situated agents, and the future of cognitive science'", *Frontiers in Psychology*, 3, pp. 1-2.

This assonance between Predictive Processing and the concept of AOR solicits me to formulate the following *working hypothesis* about Clark's more recent work: Clark's Action-Oriented Predictive Process is a neuroscientifically based and implemented version of his AORs. My general idea is the following one: in Clark's 1997 book, the brain fitted with his picture of AORs just because action-oriented representations were identified with neural populations. Now, anticipatory and control aspects of the brain, which were at the core of Clark's concept of AOR, are taken into account in great detail.

The philosophical point that this working hypothesis entails is this one. If a test of my hypothesis will demonstrate that Clark's Action-Oriented Predictive Processing still sticks to the problem of AORs, then Clark's project of building a "mature science of the embodied mind" (Clark 2015b: 16), which lies on the same ground of so-called enactive accounts (Clark 2016: 235, 291), is not successful.

This idea of pulling out again the philosophical worry about representations is not only motivated by the *fil rouge* of my research -which draws a comparison between the EMH and the enactive approach focusing on the problem of representations- but it is also connected to the very reason why Clark talks about Action-Oriented Predictive Processing. Indeed, Clark clearly states that Action-Oriented PP "sets the scene for peace to be declared between the once-warring camps of *representationalism* and *enactivism*" (Clark 2015a: 2, emphasis added).

Considering what Madary notices in his article "Extending the Explanandum for Predictive Processing" (Madary 2015), Clark says that enactivist approaches to cognition and Action-Oriented PP are engaged in the same project: "depicting the organism and the organism-salient world as bound together in a process of mutual specification" (Clark 2015a: 3). Clark's point is the following one: since Action-Oriented PP and enactivism are engaged in the same project, and since Action-Oriented PP does not make use of *classical* versions of representational mechanisms and internal models rejected by the enactivists (see also Williams 2017), we can set the ground for an integrated science of the mind that finally puts an end to "representations wars" (Clark 2015a). Therefore, to understand if the peace treaty signed by Action-Oriented PP is effective or not, one should investigate Clark's proposal in detail, by paying attention to concepts that recall internal representations, in particular AORs.

As previously said, the huge theoretical similarities between Action-Oriented PP and Clark's previous work lead me to suspect that AORs are still there. The fact that in *Surfing Uncertainty* Clark does not give space to an explicit and detailed discussion of a possible theoretical relation between AORs and his new theory does not guarantee that AORs disappeared. As Daniel Hutto says in his article "Exorcising action oriented representations: ridding cognitive science of its Nazgûl" (Hutto 2013), "AORs are not easy to kill off. [...] Like the Nazgûl – the undead monsters of Tolkien's *The Lord of the Rings* – they (and their contents) are able to survive by finding new shapes to wear and new vehicles to ride" (Hutto 2013: 145).

III.3 An internalistic version of Predictive Coding: Hohwy's proposal.

In this paragraph, I will consider some of the core ideas of Predictive Coding. This discussion is meant to be a preparatory stage for what will be done in the following paragraphs, where I will discuss Andy Clark's approach to Predictive Coding, and enactive approaches to that model.

Here I do not want to offer a complete and detailed discussion of Predictive Coding models of the brain[106]. On the contrary, my aim is more modest. I will focus on the broad idea of the mind that this approach to the brain suggests, and I will discuss some core concepts of this model, philosophical reassessments of this neuroscientific model make use of in their explanations.

Predictive Coding -whose theoretical roots can be individuated in German physiologist von Herman Helmholtz's work on perception (Helmholtz 1866; Friston 2012; Hohwy 2013: 5) - holds that the brain is an *inferential machine*. The core idea of this approach is that perception is anchored to the world by means of inference -causal inference in particular.

To explain this Hohwy says:

> "States of affairs in the world have effects on the brain— objects and processes in the world are the causes of the sensory input. The problem of perception is the problem of using the effects—that is, the sensory data that is all the brain has access to—to figure out the causes. It is then a problem of causal inference for the brain, analogous in many respects to our everyday reasoning about cause and effect, and to scientific methods of causal inference." (Hohwy 2013: 15)

As the citation shows, the idea promoted by Predictive Coding is that agents' interaction with the world is guaranteed by the logical job done by the brain.

At the core of this approach to perception there is a particular conception of the brain, which entails metaphysical consequences about the mind. Hohwy's Predictive Coding model argues in favor of a "neuroscientific seclusion" (Clark 2017: 2): the brain is secluded from the body and the world (Hohwy 2014: 7). To put it in philosophical terms, this means that the world

[106] For a more detailed review and explanation of these models see Bruineberg, Kiverstein, Rietveld 2016; Friston 2010; 2011a; 2011b; 2012; Gallagher, Allen 2016; Gallagher 2017; Hohwy 2012; 2013; 2014; Macpherson 2017; Menary 2015; Rao, Ballard 1998; Spratling 2015).

and the body are not "always and already there" for the brain-mind's cognitive mechanisms, but, on the contrary, they lay "beyond the veil of sensory inputs" (Hohwy 2013: 238). Sensory inputs, instead of being conceived to be an *open*[107] point of contact for the brain and the environment, are rather described as a veil that divides information from its hidden causes. This veil is usually called "Markov Blanket"[108].

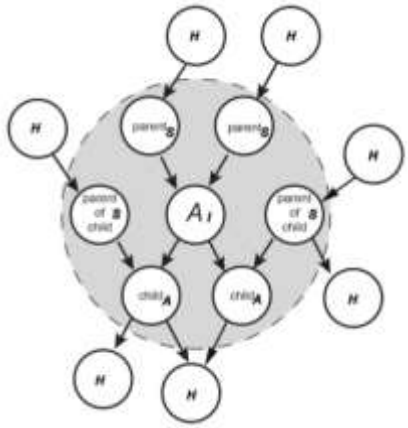
Given this separation between the brain, the body, and the world where the agent consciously perceives objects, the core idea of Hohwy's discussion is the following one. Since the brain is not in direct contact with what caused the stimuli it receives from the senses, the brain is always engaged with the task of guessing the sources of its own states. This operation of guessing -which is performed at the sub-personal level, within the boundaries of the Markov blanket (Hohwy 2017: 4)- aims at making sense of pieces of bottom-up information. To make sense of sensory information, the brain does something that has the same structure of inference, in particular *inference to the best explanation* (Hohwy 2014: 5).

The idea that lies at the core of this neuroscientific account of perception, is motivated by the observation of a general experiential fact.

[107] Notice that this is my interpretation of the concept of evidentiary veil in Hohwy's Predictive Coding. Here I claim that the evidentiary veil is like a wall that demarcates the brain and the world because I conceive of openness in a more phenomenological or enactive sense. Indeed, to me, "being opened to" refers to a kind of direct contact of the agent with her subjective world, guaranteed by the constitutive coupling of the agent's states of action and perceptual readiness with possibilities of action distributed in the world. I think that Hohwy would not agree with my interpretation of "openness". In fact, when he deals with this problem, he says that his view "does not conceive of the mind or the brain as causally insulated from the world around it. Indeed, this view must conceive of the mind and the world as causally linked, through the causal interface of the Markov blanket" (Hohwy 2017: 13). To me, this way to conceive of openness is misleading. Saying that mind and world are linked by means of a causal and inferential interface logically entails that the causes of what is inside the brain have to be inferred, then they are hidden. If objects that cause states of the brain are hidden to the brain, then world and mind are not connected by a relation of constitutive openness.

[108] This concept has been developed by the mathematician Andrey Markov, who explored abstracts systems able to remember their past trajectories only insofar they store a single current value. In these systems, the next state depends on the current state only. The notion of Markov blanket has been developed in the cognitive sciences to describe sets of nodes such as that, for a given node X, the behavior of X can be predicted by knowing the states of the other nodes. This means that to predict the state of a given node, all you need to do is to look at what is inside the Markov blanket, namely other states of the nodes of the blanket (Clark 2017: 1-2).

Figure 6: Markov Blanket
[Picture taken from Gallagher, Allen 2016]



"In our complex world, there is not a one-one relation between causes and effects, different causes can cause the same kind of effect, and the same cause can cause different kinds of effect" (Hohwy 2013: 13). This is to say that objects or features of the world can give rise to similar effects on sensory organs, and this means that, to the brain, the information conveyed by the senses can be consistent with many causes.

For example (Hohwy 2013: 13), think of the potentially identical sensory inputs caused by a bicycle or by a picture of a bicycle. The information channeled by the senses is very similar: in both cases, senses inform the brain about the color, the shape, and other features of the object perceived. That is why the brain should perform inferences to the *best explanation*. This means that it has to explain away the occurrence of some evidence (the brain own sensory state), by providing evidence by itself (Hohwy 2014: 5). It has to provide evidence for the degree of probability that the set of sensory information entertains with its hidden cause (the object of the world), which is represented by the brain as a prior belief. In this sense, the brain provides evidence by itself: the prior (namely the hypothesis about the hidden cause of the sensory stimuli) explains evidence (information channeled by the senses), and evidence, on its turn, gives an explanation to the hypothesis.

The postulation of such a mechanism of evidence to the best explanation is theoretically connected to the idea that the brain is a *hypothesis testing machine* and a *Bayesian engine*. Hohwy's idea is that, in order to explain how the brain succeeds in its job of guessing operated by means of inferences, we should individuate a kind of normative dimension. We should find something that not only explains what the brain does when it is engaged in predictive perceptual mechanisms (i.e. when it links bottom-up flows of information with top-down informational flows, sometimes called "prior beliefs"; Hohwy 2013: 15), but

that also explains what the brain *should* do in order to get the world right (Hohwy 2013: 14). To put in other words, the theory should individuate normative principles or rules that explain why some signals are computationally more relevant than others (Heilbron, Chait 2017) in the operation of getting the world right.

Hohwy finds this normative dimension of perceptual mechanisms of the brain in probability theory, in particular in Bayesian epistemology. To put it roughly, his idea is that Bayesian epistemology offers Predictive Coding insights about the rules that the brain follows when it makes its inferences to guess the hidden causes of its current states. Looking at Bayesian epistemology, Predictive Coding can explain *why* the brain "thinks" that that flow of bottom-up information is consistent with that prior belief concerning the hidden cause of sensory information and not with other priors.

To explain the role of Bayes' rule, and then to explain the idea that the brain is like a scientist who makes discoveries through hypothesis testing (Hohwy 2017: 4), Hohwy begins its argumentation considering a very simple case of everyday perception (Hohwy 2013: 15 and ff.).

An agent is in a house with no windows and no books or internet. She hears a tapping sound. To make sense of this perceptual episode, she needs to figure out what is causing that noise. The agent starts to wonder about the cause of the noise, and she makes a list of the possible causes of the input. It could be a woodpecker pecking at the wall, a branch tapping at the wall in the wind, music from her neighbor's house, and so on. It could be also that her house has been launched into space overnight and the sound is produced by a shower of meteorites. The list of possible causes is endless. Those possibilities are *hypotheses* concerning the cause of the noise: that is why Hohwy states that "the problem of perception is how the right hypothesis about the world is shaped and selected" (Hohwy 2013: 16).

The first criterion the agent follows in order to test her hypothesis and to choose the hypothesis that bests explains the perceptual episode is *likelihood*. To assess likelihoods, the agent relies on assumptions about causal regularities she observed in the world, for example typical effects of woodpeckers, her neighbor's habits, and so on. Looking at her previous experiences, the agent can rank hypotheses according to their likelihood, namely according to their closeness to the effects they should explain. Likelihood explains the way the agent judges the "goodness" of each hypothesis, and the way she ranks the *predictive* power of her hypotheses (namely how much a hypothesis is good at predicting its effects, which are what the agent gets from the senses). This explanation grounded on the concept of likelihood suggests that the problem

of perception can be simplified by making sets of hypotheses progressively smaller and smaller, giving more weight to hypotheses with a high degree of likelihood.

To Hohwy, the problem of perception in the brain follows the same logic of this everyday perceptual task. The brain is like the agent of the thought experiment, the house without windows is the agent's skull, the tapping sound is the sensory input, and the agent's previous experiences are the repeatable patterns according to which sensory inputs show up in the brain.[109] Crucially, inferences to the best explanation the agent consciously makes are like the brain's *unconscious perceptual inferences*, and the ranking of hypotheses according to their likelihood is like the brain's operation of weighting its top-down predictions (priors) according to Bayes' rule (Hohwy 2013: 18).

Basically, Bayes' theorem is a statistical law that describes the probability of an event by relying on previous knowledge that can be related to the event taken into account. More precisely, the rule tells to update the probability of each hypothesis (e.g. "the tapping sound is caused by a woodpecker") i) given some evidence (e.g. hearing some tapping sound), ii) and by considering the prior probability of the hypothesis according to criteria of likelihood, which is an estimation of how probable that hypothesis is independently from the effects currently observed (Hohwy 2013: 17).

According to Predictive Coding models, the brain embodies such a probabilistic rule because the explanation holds that the brain works by representing "the *statistical* structure of the world at different levels of abstraction", and it does so "by maintaining different causal *models* that are organized on different levels of a *hierarchy*, where each level obtains input from its subordinate level. In a feed-backward chain, predictions are made for the level below" (Blokpoel, Kwisthout, Rooij 2012: 1; emphasis added).

The concept of *hierarchy* is pivotal in this model, because it accounts for the very predictive feature of perceptual mechanisms, or unconscious perceptual inference. According to explanations offered by Predictive Coding, the models embodied by the brain are stored according to a hierarchical

[109] Notice that the metaphorical connection between the agent's previous experience and repeatable sensory patterns is not mentioned in Hohwy's 2013 book. Nevertheless, this connection seems to be motivated by what the author says in his 2012 article "Attention and conscious perception in the hypothesis testing brain", *Frontiers in Psychology*, 3, pp. 1-14, where he links repeatability of sensory inputs to the formation of expectations (Hohwy 2012: 2).

structure, which organizes statistical regularities. Looking at conscious experience one can observe that regularities come at different time scales. There are fast time-scale regularities, which include things like how shadows change as you move an object in your hands, and slower regularities, such as those which concern the trajectory of a balloon the agent is trying to catch. Those regularities can be also analyzed looking at their degree of detail. Fast-time regularities are more detailed, slower regularities are more abstract. Moreover, regularities can be organized from faster to slower, from more detailed to more abstract.

The hierarchical order of those regularities is recapitulated by the cortical structure of the brain (Hohwy 2013: 28). Faster timescale regularities are represented in low levels of the hierarchy, slower regularities are represented in higher levels. Since regularities are not only organized according to their position in a given timescale, but also according to their degree of abstraction, it can also be said that lower levels of the hierarchy represent variants of experience (e.g. as the agent's "eyes or head move or objects of perception shift around, the brain needs to process fast causal regularities for very basic sensory attributes such as contour, shading, and orientation"; Hohwy 2013: 29), higher levels represent invariant aspects of experience (e.g. the perception of objects as enduring).

The intertwinement of those levels is crucial to understand perceptual experience. Indeed, the core idea of Predictive Coding is that correct guessing is given by the right match of general expectations about the hidden causes of the signal and the variables of experience represented by the incoming signal. The hierarchical structure of this model is said to explain this phenomenon because it postulates that, in the process of perceptual inference, there is a cascade of predictions that are constantly redefined and made more "concrete" when they descend along the hierarchical structure. As a matter of fact, each level takes information from the level below as its input. This helps the higher level choose the best explanatory hypothesis, given the evidence provided by bottom-up cascades of information.

The philosophical point this hierarchical approach to information processed by the brain suggests is that perception, rather than being a bottom-up process in which the brain simply receives a rich signal conveyed by the senses, is an active or *generative* process. Indeed, the way Predictive Coding describes flows of information in the brain suggests the following story: the brain *generates its own perceptual states*. It does so by formulating hypothesis about the hidden causes of perceptual stimuli, and by testing them with sensory

evidence coming from lower levels of the cortical hierarchy. Instead of just receiving information from the senses, the brain generates information, by representing predictions at different levels of the hierarchy, and predictions are states in which the brain is at different time scales. That is why the brain is said to generate its own states.

Now, even if at first glance one can think that this model -to the extent that it gives a lot of space to concepts such as prior, belief, model, hypothesis, and, crucially emphasizes generative aspects of the brain- offers a picture of the brain (and then of cognitive processes) that gives more weight to top-down effects than to the bottom-up ones, a closer look at the explanation shows that the aim of the model is slightly different. As a matter of fact, Predictive Coding states that, in order to offer a successful explanation of cognitive mechanisms, the theory should account for the balance between bottom-up and top-down (Hohwy 2013: 67).

The way the theory accounts for this balance can be explained if we consider in more detail the function of signals coming from lower levels of the hierarchy. The crucial function of bottom-up signals is to give *feedbacks* on the internal model of the world (Hohwy 2013: 47).

Pivotal in this process of perceptual inference and tracking of the brain's own causal structure is one particular kind of feedback: *prediction error*. Prediction error is a signal that "says" that top-down predictions do not match signals from the lower level of hierarchy; it gives information about the difference between the actual current signal and the predicted one (Clark 2013a: 2).

This difference between the current signal and the predicted one is said to give rise to an effect surprise, or surprisal, to use a term that is less experientially loaded (Hohwy 2013: 52; Clark 2013a: 3). To some extent, "surprisal" is close to experiential surprise. A person experiences surprise when she finds herself is in a state she did not expect to be in, because the occurrence of that state had been judged to be not that probable in the past (Howhy 2013: 52). Similarly, when prediction error occurs, the brain finds itself in a state of surprisal (or "self-information"; Friston 2010: 2), where surprisal is the "sub-personally computed implausibility of some sensory state given a model of the world" (Clark 2013a: 6, who cites Tribus 1961). Suprisal is a declining function of probability: "as the probability goes to zero the surprisal goes up, as the probability goes to 1 the surprisal goes down" (Hohwy 2013: 51).

Surprisal and prediction error are linked to the pivotal concept of *Prediction Error Minimization* (PEM) and to that of *Free-Energy*[110]. The point that lies at the core of PEM is twofold.

First, it is theoretically connected to the idea that the brain, instead of dispersing information and energy through a large set of states (Hohwy 2017: 2), self-organizes itself in order to occupy a limited range of states, namely to maintain its homeostasis; Friston 2011b: 92). This self-organization allows the brain to protect itself from entropic disorder (Hohwy 2013: 191), and then to be in a state of equilibrium. Basically, the idea is that the brain works in order to prevent the organism from finding itself into states where the organism is not expected to be found in the long run. "For any given organism, there is a set of states where it is expected to be found, and many states in which it would be surprising to find it" (Hohwy 2015: 2).

To maintain its equilibrium, the brain has to minimize what is called "free-energy". According to information theoretical terminology, free-energy is a quantity given by a system's average energy minus its entropy or disorder (Friston, Stephan 2007: 421; Friston 2011b: 93-94; Howhy 2013: 180). The free-energy principle "says that any self- organizing system that is at equilibrium with its environment must minimize its free energy" (Friston 2010: 127). To minimize free-energy, and to guarantee itself a state of equilibrium in which the occurrence of unexpected or surprising states is minimized, the brain should find itself within its expected states. Expected states are estimated in prediction error minimization (Hohwy 2015: 4). Therefore, PEM mechanisms constitute a great source for "the life of the brain", because they allow it to find an equilibrium.

PEM has also another very important function, which concerns the epistemological points sketched out previously: it allows the brain to have a high probability to get the world right. To explain this with the concepts at the core of Predictive Coding mentioned above, it can be said that PEM gives bounds to the brain's job of guessing the hidden causes of its own states performed via perceptual inferences by creating a close fit between predictions and actual sensory inputs.

Hohwy says that there are two main ways to minimize prediction errors (Hohwy 2013: 43; Hohwy 2015: 4), which have different directions of fit.

[110] See Friston, Stephan 2007; Friston 2010 for a detailed discussion of the role of the Free-Energy principle in the brain.

First, the brain can work on its own models by revising their parameters, namely by revising hypotheses to make them more suitable to explain sensory evidence. This way of conceiving of PEM lies in the realm of perception, explained in terms of hypothesis to the best explanation, as previously said, and basically accounts for PEM in terms of optimization of internal models (Hohwy 2015: 2; 2017: 2).

Second, prediction errors can be minimized by *sampling* more carefully, namely by trying to keep the model parameters stable and using them to generate predictions. This second way to minimize prediction errors, Hohwy says (Hohwy 2015: 4), is action. Action provides proprioceptive predictions, which are delivered to the reflex arcs and fulfilled there until the expected sensory input comes.

This way to conceive of action as a behavioral fulfillment of expectation is very specific: it is compared to the job of a scientist which retains hypotheses (which correspond to the brain's prediction) by controlling the environment until the expected evidence occurs. Moreover, since prediction error minimization by means of action, in the long run, approximates Bayesian inference, action can be said to be inferential (Hohwy 2017: 2). That is why this way of conceiving of action in PEM is labeled "active inference" or "embodied inference" (Friston 2011b).

At first glance, phrases such as "active inference" and "embodied inference" seem to suggest a less internalistic view of Predictive Coding.

At the beginning of my discussion of Hohwy's Predictive Coding, I pointed out the emphasis that this model puts on neuroscientific seclusion, that is the idea that the brain is isolated from the body and the world. I said that this isolationistic conception of the brain justified the inferential nature of the model. Since the brain is isolated from the world, it has to infer the causes of its own states, which represent object or features of the world. The role given to active inference seems to challenge this isolationistic view. That is why -as it will be explained in the following paragraph- extended and enactive conceptions of Predictive Coding precisely stress on this point in order to make their theory matching the core assumptions of their philosophical proposals.

Nevertheless, this openness to the world that active inference would guarantee to the guessing brain is not accepted as a good philosophical point in Hohwy's model, which defends its internalism at many explanatory levels.

On the one hand Hohwy acknowledges that PEM in terms of active inference gives a key role to the body in perception. As a matter of fact, the way sensory inputs are sampled by means of action is shaped by the bodies we have (for example, think of an occurrence of active inference where an agent figures

out how sensory inputs will change according to the way an object is manipulated using hands). On the other hand, he says that in the cognitive mechanism of Predictive Coding the agent's body really matters only to the extent it is represented in the brain's models, as a parameter for PEM. What matters to Predictive Coding systems is not the "lived" and "motoric body" that performs actions in the world, but the functional role of the body in the economy of the parameters of PEM, parameters the brain is said to have access to through representations.

Moreover, Hohwy clearly states that both perceptual inference and active inference are performed by the brain, within the boundaries of the Markov blanket, where evidence is provided to the system. The distinction between inner and outer of the evidentiary boundary or veil is strict: inference is done by the inside of the mind, and external action just provides the brain with inputs that will fulfill its expectations represented in the model (Hohwy 2014: 7; 10; 16-17).

Therefore, the philosophical point at the core of this approach can be expressed in this way. Agents' engagements with the world are usually smooth and successful because those worldly engagements are guided by anticipatory mechanisms. Those anticipatory or predictive mechanisms, in order to work well, need some kind of evidence. Since those predictive mechanisms are internal mechanisms of the brain, and given that the brain is isolated from the body and the environment, evidence, namely what guarantees the system epistemic efficiency, is internal to the brain's mechanisms themselves (namely it is given within the boundaries of the Markov blanket). Since what actually guides the epistemic efficiency of anticipatory mechanisms is internal to the brain, and since anticipatory mechanisms of the brain guide smooth action-perception loops, then what actually explains action perception-loops is internal (brain-bounded).

III.4 Clark's approach to anticipatory mechanisms: extending Predictive Processing?

In the previous paragraph, I briefly sketched out the core ideas of the Predictive Coding theory, and I emphasized that Hohwy's version of this theory is explicitly and strongly internalistic. In this paragraph, I will consider Clark's approach to Predictive Coding, which he calls "Predictive Processing" (PP)[111] or "Radical Predictive Processing" (Clark 2017; 2016; 2015b; 2015c; 2013a; 2012). In particular, I will focus on the way he tries to make this explanatory framework matching the EMH.

To do that, I will develop my discussion of Clark's approach to Predictive Processing by taking into account four key concepts: Markov blanket(s), productive laziness, Parity Principle and Predictive Processing, and active inference.

i) Markov blanket(s)

As pointed out in the last pages, Hohwy's internalistic claim is dependent on a specific conception of the brain -which I called "neuroscientific seclusion" following one of Clark's recent articles (Clark 2017: 2)- according to which there is a clear and strict demarcation between the brain and the world. This peculiar conception of the brain -which is at the core of the justification of the inferential conception of cognitive processes- finds its place in the explanation of Predictive Coding by means of the concept of Markov blanket.

The core claim of Hohwy's Predictive Coding is that states of the sensory system determine a Markov blanket. The Markov blanket defines the boundaries of the mind. Hohwy's idea is the following one. In the same way the behavior of nodes of abstract systems can be fully predicted by knowing the states of other nodes of the system, states of the brain can be predicted checking other states of the brain. That is why the brain is said to be a self-evident engine: to discover the hidden causes of its sensory inputs, the brain just has to know information about its own states, which are the bricks of generative models. This makes what is outside of the Markov blanket (the body and the world) uninformative once the states of the blanket are known.

[111] Notice that Clark's terminological shift from "Predictive Coding" to "Predictive Processing" aims to emphasize the difference between these two views of anticipatory mechanisms. In particular, the phrase "Predictive Processing" is said to be expressive of a more dynamical view of anticipatory mechanisms. As it will be explained in this paragraph, Predictive Processing is said to be more dynamical than Predictive Coding because it emphasizes more the role of action in predictive mechanisms (Clark 2013a: 22), and it conceives of the boundaries of the predictive mind as flexible boundaries (Clark 2017: 16-17).

Such a consideration entails an interalistic and narrow conception of the mind: if the mind should be individuated where there is information interesting to explain the animal's behavior, and if the explanation takes into account information processed by the brain only, then the mind is individuated in the anticipatory machine instantiated by the brain. To put in other words, the mind, conceived as a cognitive machine, knows something about the world by providing evidence from its own hypothesis. The evidence that the mind provides by itself is evidence about its internal models: what the mind knows is a proxy of the world, and all parts of cognitive processes, since all the inferential guessing is performed by the brain, are internal. Then the mind is internal in two senses. First, it is internal because its object of knowledge (the model) is internal; second, what is cognitive (i.e. inference) in those anticipatory processes is performed within the boundaries of the mind, namely in the nodes of Markov blankets.

Since the concept of Markov blanket and the concepts it entails (above all that of epistemic evidence) are pivotal in the internalistic definition of the mind, an explanation aimed at redefining the boundaries of the predictive mind has to reconfigure the concept of Markov Blanket (Clark 2017; 2016).

The idea of reconfiguring the concept of Markov blanket is crucial in the project of extending PPs not only because of the theoretical implications that will be discussed here, but also because advocates of internalistic versions of Predictive Coding say that embodied, enacted and extended approaches to cognition and the mind "seek to obliterate the Markov blanket" (Hohwy 2017, as cited in Clark 2017: 6); therefore they cannot be consistent with the core ideas of Predictive Coding. A plausible way to conceive of the Markov blanket in a more extensive way would be a step towards classic views of Predictive Coding, without entailing internalistic consequences.

Clark's core idea is that one can think of extended predictive minds preserving the concept of Markov blanket if this one is considered in its mutability and multiplicity. Contrary to Howhy's position, Clark claims that we should not think of an individual, self-evidencing system bounded by an unchanging Markov blanket. Rather, Clark's idea is that PP is bounded by many Markov blankets that change over time.

The theoretical point that lies at the core of this way of conceiving of the Markov Blanket is the idea that the Markov blankets boundaries we choose depend on the explanatory purposes of our theory. For example, if the aim of an explanation is to consider a system stable over time, bounding the system by

an individual, unchanging Markov blanket offers a theory that is explanatory enough.

The problem is that what Predictive Processing seeks to explain is not a stable and unchanging system. This theory aims to explain predictive mechanisms that guide animals' action-perception loops. Animals are beings that change over time. For instance, think of the human ontogenetic development. Not only bodies change over time, but also, and crucially, human beings are "mentally metamorphic" (Clark 2017: 13). This suggests that the boundaries within which cognitive processes (i.e. Predictive Processing) take place change over time, and this motivates the idea of thinking of the Markov blanket as a fluid or changing veil, which expands its boundaries in tandem with mental development[112].

Moreover, Clark suggests that mental development is fostered by the use of sets of tools, strategies and devices (neural, bodily and extra-somatic). The use of such devices not only produces dramatic alterations in the human brain[113] (and this suggests the idea of a changing Markov blanket), but it also produces changes in agents' cognitive experiences over time. The more an agent becomes skillful in her use and manipulation of such devices, the more her set of cognitive possibilities will extend.

As already explained, the EMH explains this augmentation and improvement of cognitive possibilities by appealing to the concept of coupling, which describes the human agent and cognitive external tools as a unitary system. The notion of organism-environment system challenges the idea that cognitive activities are bounded by an individual Markov blanket. Indeed, in Clark's EMH, coupled systems are created every time an agent engages with an external cognitive device that allows the cognitive load of a task to be distributed between the biological agent and an extra-somatic resource. This suggests that wherever new technologies interface with bio-systems a new Markov blanket can be individuated. For instance, think of an agent using Google-glass (Clark 2017: 12). The agent plus Google-glass is a new cognitive system, which performs cognitive activities within a certain kind of bound (that one defined by the coupled system). This does not mean that "old" Markov blankets, such

[112] Here I consider an example that concerns human knowledge, but other striking examples concerning non-human animals can be found. For example, think of animals such as metamorphic insects, which undergo dramatic changes during their metamorphoses. Those dramatic changes are examples useful to question the idea that the Markov blanket is static, since metamorphoses change the terms of living systems, while living systems still maintain their life-unity (Clark 2017: 11-12).

[113] For example, think of Clark's discussion about the structural changes that language development and use produce in the brain (Clark 1998).

as that one which bounds the self-sustaining activity of cellular organelles, are erased. Markov blankets produced in experiences of coupling constitute layers of self-maintaining processes. The demarcation of each Markov blanket explains each self-sustaining process, by individuating the boundaries of interest that define a cognitive practice performed at that time t.

This idea of flexible and multiple Markov blankets allows the theory to think of a *porous* demarcation between the predictive mind and the environment (Clark 2016, in particular Chapter 8), challenging the isolationistic view of the brain offered by Hohwy, and providing the theoretical ground for a theory of the Extended Predictive Mind. At the same time, Clark's approach to Markov blankets given in his 2017 article, meets the philosophical requirements of classical views of Predictive Coding, which seek to find a clear definition of the boundaries of the predictive mind. As a matter of fact, Clark's multiplication of Markov blankets, which works in tandem with the idea of coupling, guarantees a closure to cognitive systems, and closure is precisely what classical approaches to Predictive Coding look for.

The important philosophical point of Clark's closures consists in the way Markov blankets are individuated, namely in the idea of thinking of Markov blankets as the boundaries of extended self-sustaining systems constituted by the agent and external cognitive devices, which shape and structure cognitive practices.

ii) Productive laziness

Clark's idea of redefining the boundaries of the predictive mind by thinking of a multiplicity of Markov blankets that define coupled predictive systems is supported by the consideration of a recurrent theme in works in embodied and environmental situated cognition: the value of the brain's "productive laziness" (Clark 2015c: 9-12; 2016: 244- 245).

As pointed out in the previous paragraph, Hohwy's theory of Predictive Coding is focused on a detailed account of the predictive powers of the brain, explained through the concepts of generative model, hypothesis formation and hypothesis testing, and self-evidence in processes of inferences to the best explanation. This account of Predictive Coding describes the brain as an intrinsically active engine, which has all the means for its self-sustainment and for deploying veridical cognition inside of itself.

The picture of the brain that Clark's PP offers is similar, because Clark too, faithful to the core ideas of Predictive Coding, puts strong emphasis on active powers of the brain, stressing on the idea of models that are self-generated and

that actively shape perceptual data by means of top-down predictions (Clark 2013a; 2015b; 2016).

Nevertheless, by taking into account "productive laziness", Clark's theory points to an explanation of the active, productive, and pro-active brain, which is said to be consistent with embodied and extended views of the mind.

"Productive laziness" is an expression that Clark borrows from Herbert Simon (Simon 1956; Clark 2015c: 9), and that has been coined to explain economical but effective strategies in fast-time problem-solving tasks. An example that explains productive laziness in cognitive tasks is the following one (Clark 2015c: 9). An agent has to choose a restaurant for a special occasion. Instead of reading all the reviews and menus for every restaurant within a 5 mile radius, the agent chooses a restaurant that a trustworthy friend suggested her. The cognitive strategy the agent chooses is the faster and simpler one: it is a solution that allows her to waste a minimum amount of energy and time, but at the same time, it is a successful heuristic, because it actually helps the agent get the result expected.

This idea of productive laziness can be extended to many cognitive practices, and it can be applied to an explanation of how the brain works. Indeed, the idea the concept of "productive laziness" suggests is that the brain mandates parts of its job to the body and to environmental props.

A core strategy for the distribution of the cognitive load of a task between brain, body, and world is the ecologically efficient use of senses in action.

To explain this point, which can be described in terms of "ecological balance" (Clark 2016: 245), and that theoretically goes back to general ideas of ecological psychology, Clark considers the outfielder's example (Clark 2015c: 11; 2016: 190, 247, 256). The example is about a baseball player who should run so as to catch a "fly ball" in baseball. Old-fashioned explanations of perception would try to account for this cognitive episode talking about the complex job done by the visual system, engaged in the transduction of information about the current position of the ball. The outcome of this job of transduction is said to be an internal representation of the future trajectory of the ball. This internal representation would allow the agent to be successful in her practice because it would provide the internal reasoning system with the cognitive means to predict its future perceptual states.

One of the problems of this explanation is the complex and theoretically obscure internal system it postulates in order to account for cognitive efficiency. A different explanation can be formulated looking at the way agents, during their conscious experience, structure their perceptual fields in order to make predictions about future perceptual states faster and less cognitively

demanding. This environmental structuration is provided by action, seen as a fast and cheap heuristic. According to this explanation, the baseball player, instead of accurately representing sensory stimuli in the brain, and transmitting these data to another distinct reasoning system, directly acts upon the world: she runs in a way that seems to keep the ball moving at a constant speed through the perceptual field. The movements performed by the player predict the future position of the ball (and the agent's sensory states as well) because the outcome of the agent's own movement is a cancelation of apparent changes in the ball's optical acceleration: this will allow the agent to end up in the point of the field where the ball will hit the ground.

This simple example -which is just one of the many cases in which the lazy brain mandates its cognitive job to the body- is crucial to understand Clark's account of PP because it emphasizes two key features of Clark's reassessment of classical Predictive Coding account.

First, the football player's case illustrates the core claim of Clark's PP: "real-world prediction is all about [...] selection and control of world-engaging action". That is why the predictive machines Clark talks about are "Frugal Action-Oriented Machines", aimed at "helping animals achieve their goals while avoiding fatally surprising encounters with the world" (Clark 2016: 250). This point is particularly important for the consideration of the philosophical consequences of Clark's PP. Indeed, on the one hand, the very strong emphasis that Clark puts on predictive mechanisms in action-control and action-selection is a theoretical shift from classical views of Predictive Coding. In fact, even if also classical views deal with action, those explanations are more focused on hypothesis-testing operations performed by the brain. This makes them really internalist accounts of predictive mechanisms. On the other hand, Clark's interest in "Frugal Action-Oriented Machines" confirms the working hypothesis formulated in §III.2, namely the idea that Clark's PP and the concept of AORs developed in *Being there* seek to explain the same phenomenon: on-line, context-sensitive action control and action selection.

Second, the football player's example, and the way it accounts for the brain's "productive laziness", is suitable to introduce another key concept of Clark's PP: Parity Principle in the explanation of predictive mind. Indeed, it seems that the story about productive laziness previously described -which says that, since the brain is productive but lazy, it mandates parts of its cognitive load to the body and the world, adopting "mix and match strategies" (Clark 2016: 252)- seems to be just a different way to express the idea of cognitive impartiality.

iii) Parity Principle

Even if when Clark discusses PP he does not mention the Parity Principle (Clark, Chalmers 1998) explicitly, implicit occurrences of this this concept are frequent, especially when he deals with the broad philosophical implications of PP (Clark 2016).

For example, he writes:

> "Actions that engage and exploit specific external resources will now be selected in just the *same* manner as the inner coalitions of neural resources themselves.[...] Invoking a bio-external resource, and moving our own effectors and sensors to yield high-quality task relevant information are here expressions of the *same* underlying strategy, reflecting our brain's best estimates of where and when reliable task-relevant information is available" (Clark 2016: 260, italics added).

This citation has two core points. First, it suggests that the cognitive system selects external and internal resources in the same manner. This means that resources the system relies on for its predictions are not selected according to their spatial location or their material substrate. Second, Clark says that bio-external resources, sensorimotor activities, and neural resources are part of the same cognitive strategy; they have this constitutive role in the cognitive system because both channel information that is relevant in the economy of the cognitive task. These two points clearly recall the Parity Principle, pivotal concept of Clark's extended functionalism. Indeed, what seems to allow extra-neural and extra-somatic resources to be part of the predictive system is the relation of similarity their functions entertain with brain functions, and this relation of similarity is individuated in their capacity of providing informational contents to the system (Kirchhoff 2015: 3).

Even if at first glance the introduction of the Parity Principle in the PP framework can sound like an appealing strategy to extend PP, I claim that this is not a good philosophical proposal. As already discussed in §I.9, the Parity Principle is a tricky concept. Indeed, it implicitly entails an internalistic prejudice, because it allows extra-neural resources to be part of the cognitive system because they are similar to internal processes, which are considered to be the first term of the comparison that the principle draws. The Parity Principle, by giving a cognitive primacy to neural resources –primacy that depends on conceiving of internal resources as models for external cognitive processes- implicitly puts the cognitive specificity of extra-neural resources in the background. Therefore, since the Parity Principle is implicitly applied to

the PP framework, the "extended" explanation of anticipatory processes seems to suffer from the EMH's disease, which I would call "concealed internalism".

Moreover, the informational conception of cognitive processes that the Parity Principle entails is a challenge for a fully embodied account of PP. If extra-neural resources play an interesting role in anticipatory cognitive mechanisms only because they function as information carriers, what about the role of the actual "sensorimotor body", the body that does things in the environment? And also: what does the job of information processing? The body or the brain?

Saying that the body is crucial in predictive processes because it carries information relevant to the system is different from claiming that the body, thanks to its situated actions, is crucial for the process of prediction. In order to predict peace between extended and enactivist approaches to cognition (Clark 2015b: 16), PP should account for this aspect of the embodiment as well. Moreover, if it was the case that information carried by the body was said to be processed by the brain only, then the explanation would give the body an important, but still secondary role in PP, because this would entail that core control functions would be individuated in neural resources.

To explore these issues, now I consider the main way Clark accounts for the role of the body in his PP framework: active inference.

iv) Active inference

The concept of active inference has been already introduced implicitly in my discussion when I considered Clark's outfielder's example, aimed at explaining the contribution offered by embodied action to predictive mechanisms. Indeed, to Clark, action coincides with active inference, namely the combined mechanisms through which perceptual and motor systems work together to reduce prediction errors (Clark 2016: 122).

The idea that lies at the core of active inference is that we can think of the brain as always engaged in finding neural states that best accommodate incoming sensory signal. The concept of active inference accounts for this capacity of the brain in terms of world-engaging action (Gallagher 2017: 17-18): when the agent moves her body in a certain way, she generates the sensory consequences expected by the brain (Clark 2013a: 6). This helps the lazy brain to minimize the cognitive load it has to deal with.

One of the problems of Clark's account of active inference is precisely connected with the distribution of the cognitive load active inference is said to allow. Indeed, even if Clark actually emphasizes the very role of embodied action, describing cognitive episodes in which agents actually do something in

the world in order to foster prediction, when he has to account for the way information provided by the senses is processed, he seems to ascribe this job of information processing to the brain, conceived as the center where information is manipulated (see Gallagher 2017: 18 for this critique). Indeed, Clark says that "action serves to deliver fragments of information 'just in time' for use, and that information guides action" (Clark 2016: 250); later he says that action control is guaranteed by the brain's continuous computations of possibilities for action (Clark 2016: 251).

The problem of this explanation obviously does not rest in the fact that Clark gives importance to the job of the brain -also enactivists widely consider the role of the brain in cognitive processes-, but rather lies in the centrality ascribed to brain, and to the concept of computation. This seems to be a problematic point in Clark's project of "putting predictive processing, body, and world together again" (Clark 2015b: 11). If one the one hand it is true that Clark's PP actually gives theoretical space to the world and the body, on the other hand it is also true that the body and the world play just a vicarious role in the explanation. It is not the real, material world or the anatomical, sensorimotor and lived body that are crucial for PP. What is central in the explanation is information that what is outside of the brain conveys, and the computational job the brain does on this information.

This centrality of the brain's computational abilities justifies Clark's conception of action as inference -namely the idea that action provides the internal system with information crucial for its inferences-, and it is also connected to another misleading concept (see Gallagher 2017: 19 for this critique) crucial in Clark's account of active inference: sampling.

When Clark takes into account active inference he describes the outcome of action as a sample of the world (Clark 2016: 121; 251). To explain this point Clark writes that

> "the simplest way in which a PP-style organism might be said to actively construct its world is by sampling. Action here serves perception by moving the body and sense-organs around in ways that aim to 'serve up' predicted patterns of stimulation. [...] The agent, by exposing herself to the varied stimulations predicted by the generative model, actively contributes to the world as sampled. Since it is only the world as sampled that the model needs to accommodate and explain, this delivers a very real sense in which (subject to the overarching constraint of structural self-maintenance i.e. persistence and survival) we do indeed build or '*enact*' our individual [...] worlds." (Clark 2015c: 19, emphasis added)

If I understand well what Clark claims here, action is said to be crucial in predictive mechanisms because it exposes the agent to stimulations predicted by the internal generative model and, in doing so, it contributes to the world as sampled. To Clark, the operation of sampling action is engaged in is pivotal to understand the way agents bring forth (enact) their world: by acting in a certain way, agents sample their world in a way that is suitable for their cognitive expectations.

At first glance, this explanation sounds promising, in particular in relation to Clark's project of a mature embodied cognitive science that integrates computational accounts with the enactivist approaches to cognition. Nonetheless, I suspect that this first impression is misleading. This because saying that agents enact their world in the appropriate enactivist sense entails a conception of enaction that is fully non-representational (Varela, Thompson, Rosch 1991). On the contrary, Clark's way of conceiving of enaction here seems to hide a representational commitment (Clark 2016: 293), and this precisely because of the concept of sampling.

Clark's point consists in claiming that generative models should accommodate the world sampled by action, and they do so because they are endowed with representations that have the epistemic function of "sampling the world in ways designed to test our hypotheses" (Clark 2016: 251). This suggests that, in the PP framework, action is crucial because it offers the brain samples of the world that become contents of the brain's representational mechanisms. To put in other words, what is crucial in this explanation is not the very cognitive effect that action produces in the world, structuring it as a space where cognitive practices unfold smoothly, but rather a sample of the environment that the internal PP mechanism represents in order to control action (Clark 2016: 251). What matters is not action in itself (namely what the agent actually does in the external world) but the product of action, which Clark considers to be an internal representation of the sampled world.

The fact that this representation is said not bearing richly reconstructive contents, but to "spot the context in which some frugal, action-involving procedure will work" (Clark 2016: 191), seems not being able to rescue Clark's representation of the sampled world from some philosophical problems.

To me, this way of conceiving of the epistemic role of action in terms of sampling, and not as a "doing", as an "enactive adjustment" (Gallagher 2017: 19), is unproductive if considered in the broad economy and aim of a "radical PP". Indeed, it seems to hide the same problems considered in my previous chapter, namely a weak account of the epistemic role of action. This is not acceptable in the economy of an enactive approach to cognition, which gives a

cognitive primacy to world-engaging action. Moreover, Clark's explanation, by drawing a systematical connection between action and the brain's mechanisms, ends up with depriving embodied action from its core features: being a process done or performed by an agent in the environment, thanks to the possibilities offered by her body. When an explanation circumscribes the epistemic power of action to what it offers to the brain's mechanisms, it is not actually taking into account action itself, but rather internal mechanisms that undergo action, and that have the function of controlling embodied action and perception.

These problems will be taken into account in the following paragraph, where I will offer some considerations about the role of representations in Clark's PP, in order to understand whether his account is actually suitable to extend PP or if it still sticks to an internalistic description of anticipatory processes. Moreover, this excursus about the concept of action and representation in Clark's PP is useful to consider the relation between Clark's PP and enactive accounts of predictive mechanisms. As a matter of fact, some advocates of the enactive approach to cognition, although they are critical of some concepts Clark makes use of in his PP (i.e. action as active inference, action as sampling, PP as entailing informational content; e.g. Gallagher 2017; Hutto 2017), they somehow seem to find good philosophical points for a non-representational view of PP in Clark's proposal.

For example, Gallagher emphasizes one of Clark's philosophical statements (Gallagher 2017: 17), in which Clark sounds open to the *possibility* to think of the PP's story in entirely non-representational terms (Clark 2016: 293). This, Gallagher says, would allow Clark's explanation to avoid not only rich internal representations, but also minimal forms of representationalism, such as AORs[114]. I suspect that a closer look at Clark's proposal will be demonstrative of the opposite theoretical situation: PP is a new way to rethink of AORs.

[114] Notice that at this point of his argumentation, Gallagher seems to accommodate Clark's account of PP. Nevertheless, at other stages of his argumentation, he seems to point to the direction of my working hypothesis, according to which Clark's PP hides AORs. Indeed, he notices that Clark goes back to the notion of pragmatic representation, in particular when he talks about active inference (Gallagher 2017: 19, footnote 4).

III. 5 Clark's Predictive Processing: representational flashbacks.

In the previous paragraph, I begun to sketch out of the problem of representations in Clark's approach to PP. I pointed out how representations are introduced in the PP framework, and I suggested that the representational talk is "furtive", at least if the explanation is taken into account at a broad theoretical level (i.e. by considering general statements about the nature of the predictive mind).

This idea of considering representations as epistemic posits that get into the PP framework in an unclear way is motivated by a "textual fact". Indeed, as noticed by Bruineberg, Kiverstein and Rietveld (Bruineberg, Kiverstein, Rietveld: 2016, footnote 29), there are many passages in which Clark *seems* to talk about a process that is non-representational.

For example, he states that

> "we should resist the claim that what we perceive is best understood as a kind of hypothesis, model or virtual reality. The temptation to think so [...] rests on two mistakes. The first mistake is to conceive of inference-based routes to adaptive response as introducing a kind of representational veil between the agent and the world. [...] The second mistake is a failure to take sufficient account of the role of action" (Clark 2016: 170)

Or again, he says that "the PP story does not depict perception as a process of building a representation of the external world at all" (Clark 2015a: 2).

To me, statements like those ones sound like a good starting point to draw a picture of the predictive mind that no longer relies on representations. Indeed, what Clark says seems to suggest a non-representational story in two senses.

First, the brain is said not perceiving its own internal models (or representations of the world, I would say): this seems to suggest that there can be a non-representational account of the relation between the brain and the world (i.e. it could be a relation of mere co-variance, or of causality; Orlandi 2016; Ramsey 2007).

Second, inference (in particular active inference, since Clark refers to inference when he talks about "routes of adaptive response"; Clark 2016: 170) is said not creating a representational veil between the agent and the world. As said in the previous paragraph, I suspect that we cannot avoid representations if we endorse an inferential account of cognition like that one offered by Clark.

As a matter of fact, his conception of inference has been shown to entail a representational commitment, because inference is said to be performed on representational *contents* processed by the brain. Nevertheless, considering what Clark says in a literal sense, it can still be said that he seems to promote a non-representational proposal, by saying explicitly that the evidentiary veil is non-representational.

The problem of Clark's PP framework is that it cannot fulfill these non-representational expectations. Indeed, Clark actually makes huge use of a representational vocabulary, and he does so at many explanatory levels.

First, when he speculates about the relation between the brain's Predictive Processing and the world, he emphasizes that hierarchical models are not aimed at drawing a relation of mapping between environmental and inner states; rather, they aimed at inferring the nature of the source of the signal from variations of the input signal itself. Later, when he deals with the Bayesian structure of these inferences, he appeals to the concept of *probabilistic representation*[115] (Clark 2013a: 3).

Probabilistic representations are neural representations that encode probability density distributions that are part of probabilistic generative models (Clark 2013a: 6; 2016: 39). This makes them different from traditional internal representations because, instead of merely representing given features of the world, they represent the amount of uncertainty of an internal model. To put in other words, they are said not mirroring the external world because their representational content is not about a state of affairs of the world, but it is about a probability density. For example, those representations do not simply represent "CAT ON MAT", but they reflect the relative possibility of this state of affairs (Clark 2016: 41). According to Clark's PP framework, this would make those representations suitable to overcome one of the problems classical notions of representation entail (see §I.2): the idea that our minds, because they are in contact with the world by means of propositional or picture-like representations, are like "mirrors of nature". This idea of the mind as a "mirror of nature" (Rorty 1979; Clark 1997a: 47; Clark 1995) is problematic, because it presupposes a strict demarcation between mind and world, which is not consistent with the EMH.

Moreover, Clark's account of the relation between the mind and the world in his PP makes use of a representational vocabulary when it gets to consider the epistemic strategies that the brain deploys to get the world right: PEM. Here (Clark 2016: 39) he talks about error signals in terms of "error neuros" and he

[115] For a critical discussion of the probabilistic nature of these representations see Nanay 2012.

says that those ones can be considered to be a variety of *representation neurons,* which encode sensory information that has not been explained yet.

Those two points, namely the concept of probabilistic representation and that of error representation neurons, offer a representational description of the core structure of Predictive Processing. Obviously, one can say that this claim about the representational nature of Clark's PP mechanisms should be taken cautiously, because when he deals with probabilistic representations and error representations he is offering an exam of the literature about Predictive Coding. This might entail that his own proposal is not representational, or "less representational" than those approaches. Nevertheless, it seems to me that here Clark is not making use of a representational vocabulary just in order to be faithful to Predictive Coding accounts along his excursus about those models. In fact, after this explanation of the core ideas of Predictive Coding, he considers a commentary to Rao and Ballard's seminal work on Predictive Coding (Rao, Ballard 1998) by Koch and Poggio (Koch, Poggio 1999), where the two authors express their worries about a non-representational turn in Predictive Processing. Koch and Poggio say that, since well-predicted elements of the signal are explained away within the predictive system, it seems that the explanation substitutes representations with silence. Clark answers to this anti-representational worry saying that it is unjustified (Clark 2016: 48), and he focuses on the "more and more complex representations used in processing", and he emphasizes the ubiquity of the representational information flow in the system.

This seems to be a good confirmation of a representational understanding of Clark's PP.

The problem with this interpretation of PP not only consists in the use of a representational vocabulary to denote neural populations (problem taken into account in Chapter II of this work, discussing the "standing-for" and "interpretability" conditions), but it also rests in an implicit claim according to which, since PP relies on a Bayesian account of perception, then the processes it accounts for are *fully* representational (Orlandi 2016).

As Orlandi points out, it is not clear why we should call probability density distributions (i.e. likelihoods) "representations", and why we should conceive of errors signals as representations. This for many reasons.

First, both notions of representation do not do the job representations in cognitive science are supposed to do: guiding action (Clark, Toribio 1994; Clark 2008; Clowes, Mendonça 2016; Haugeland 1991; Wheeler 2005).

As Orlandi notices (Orlandi 2016: 346), what is called "probabilistic representation" in PP -unlike in the Helmholtzian account of perception- does not serve as a premise for visual inference, but it simply marks the probability of hypotheses. Therefore, if we think of perception as guiding action, and if "probabilistic representations" do not play the role of premises in perceptual inferences, then they play no role in action-guidance.

The same thing can be said about error signals. Error signals tell the brain whether a current hypothesis is plausible or not (Orlandi 2016: 344). Unless we consider action in a very narrow and specific sense (namely as error reduction; Gładziejewski 2016), what is called "error representation" does not sound like a good candidate to guide action. Error signals just have advantages (or an "adaptive function" in a broad sense) in the inner economy of the brain, because they basically ask the brain to adjust its own states, in order "to reach an error-free equilibrium" (Orlandi 2016: 345).

Moreover, both probabilistic representations and error representations do not fulfil all the representational criteria endorsed by philosophers of cognitive sciences to determine whether an item or process X is genuinely representational or not[116].

Most cognitive scientists state that, in order to be considered to be a representation, an item or process X

i) should bear a content (= "standing-for" + "interpretability" conditions);
ii) should have accuracy conditions;
iii) should be able to misrepresent (= expressive of the "standing-for" + "interpretability" conditions);
iv) should be decoupleable from what it represent (Haugeland 1991; Orlandi 2016: 346; Gallagher 2008a; Hutto, Myin 2013; Rowlands 2006a; 2012).

[116] The list of representational criteria considered here is slightly different from the list offered in Chapter II. For example, here I do not consider the "passivity", "internality", and "genuine duration" conditions, and I introduce the "accuracy" condition, taken into account implicitly in Chapter II when I considered Hutto's objections to AORs focused on the problem of content (§II.5). I made this choice because here I follow Orlandi's argumentative structure, even if I develop some points independently (e.g. I deal with the relation between content, misrepresentation, and accuracy more explicitly).

Likelihoods, namely what has been called "probabilistic representations", do not fulfill these criteria. Indeed, they do not bear contentful states[117] (even if Clark's explanation suggests that those states are contentful; Clark 2016: 6; 21), but they rather *measure* a probability density. This entails that they do not fulfill the misrepresentation and accuracy conditions, because i) what can be inaccurate or accurate is content, ii) misrepresentation is given when the content of an item or process X does not match what it is the case in the external world.

The application of the three criteria mentioned above to "error representations" shows problems in this representational conception of error signals as well, even if in a slightly different sense. Indeed, "error representations" seem to have an informational content ("there is a mismatch between high-level predictions and the incoming sensory signal") that causes the brain to adjust its own states. Nevertheless, it is not clear how this content is able to misrepresent something and to be inaccurate. Indeed, error signals occurs when there is a mismatch between high-level predictions and the incoming sensory signal. It is not clear how error signals can misrepresent this mismatch, by giving false information about this state of affairs of the brain. Mismatches just occur, and when they occur "error neurons" are activated: this means there is a direct relation of causality between mismatches and error signals. Since those neurons do not fire when mismatches do not occur, then they cannot give the brain false information: they do not misrepresent. Therefore, also "error representations" do not fulfill the representational criteria taken into account.

The last representational condition, namely "decouplability" or "detachability" –like Orlandi calls it- needs to be considered more carefully.

In Chapter II, I considered one sense in which this concept is understood. There I defined decouplability as the possibility of an item or process X to guide cognitive processes off-line, namely when what is represented is not present in the context of the cognitive practice.

If this sense of "decouplability" is taken into account, it seems possible to claim that "probabilistic representations" partially[118] fulfill this condition, because they are used in the agent's off-line processes. The same point seems to apply to "error signals", which occur when the agent's full conscious experience of perception has not taken place yet.

---

[117] For more objections against the description of the brain's PP as contentful processes see Hutto 2017.

[118] Here I say "partially" because, as already said, "probabilistic representations" do not guide action.

The problem with this generous application of the decouplability condition rests in the fact that it entails a misleading interpretation of the context those items should be detachable from. In fact this interpretation of "decouplability" considers being off-line in relation to the agent's conscious experience. This interpretation is not actually accurate in PP, because if we look at the so-called contents of those representations, we can see that those items are not actually detached from what they are said to represent[119]. Neither "probabilistic representations" nor "error representations" are about something the cognitive agent consciously experiences. Rather, they are about the brain's internal states and models. Therefore, it is not clear why we should consider those items as detached from what they represent or as "being used off-line". If we consider the very context in which the process takes place, the neural one, it seems possible to claim that those items are used by the brain on-line, because what those items are said to represent is not absent when the brain does its job. Error signals occur when something happens in the brain (mismatches between top-down and bottom-up flows of information), and "probabilistic representations" concern something (probabilistic generative models) that is there (the model), in the brain, when the cognitive process occurs.

Therefore, given that "error signals" and "probabilistic representations" do not fulfill the representational criteria listed above, the use of a representational vocabulary to account for those core concepts of PP is unwarranted.

I claim that analogous points apply to another way to conceive of representations in Clark's PP[120], namely representations that I previously said being a reassessment of Clark's previous work on AORs (Clark 1997a; Clark, Grush 1999; Clark, Toribio 1994).

To introduce this way of conceiving of representations, I consider a passage from Clark's article *Radical Predictive Processing* (Clark 2015c), where he explains why he thinks that his conception of representations in PPs can predict peace between enactive and extended approaches to cognition.

---

[119] This is a second way to conceive of "decouplability", which Orlandi somehow considers independent from the first one (Orlandi 2016: 343). To me these two notions are interdependent, because they both refer to the possibility of an item X to play a cognitive function when what the so-called representation is about is absent in the context of the cognitive process.

[120] Here I talk about another way to conceive of representations (different from probabilistic representations and error representations) because Clark himself seems to distinguish between representations considered till this moment and representations of information conveyed by action. Nevertheless, this distinction, in Clark 2016, is not always clear. Indeed sometimes he conflates "action representation" with "probabilistic representation" (Clark 2016: 133).

> "PP [...] deals extensively with internal models [...], whose role is to control action by predicting complex plays of sensory data. This, the enactivist might fear, is where our promising story about neural processing breaks bad. [...] [Nevertheless,] the remaining distance between PP and the enactivist may not be as great as that bald opposition suggest. [...] PP, although it openly trades in talks of inner models and representations, invokes [...] representations that are fundamentally in the business of serving up actions within the context of rolling sensorimotor cycles. Such representations aim to *engage* the world [...] and they are firmly rooted in the patterns of organism-environment interaction" (Clark 2015c: 20-21; text into brackets added)

As the citation shows, Clark's appeal to internal representations in Action-Oriented Predictive Processing depends on the same *explanatory* need AORs were supposed to meet in Clark's previous work on action: serving up action in the context of rolling sensorimotor cycles, in order to allow the agent to active engage the world.

Moreover, the *nature* of those representations, I claim, is similar to that of AORs described in Chapter II. Indeed, to describe the structure of the epistemic entities considered in the citation above, Clark cites (Clark 2015c: 21) Lauwereyns' book *Brain and the Gaze: On the Active Boundaries of Vision*. In the passage quoted by Clark, Lauwereyns says that those representations are not "actual re-presentations or duplicates of objects in the world but [...] incomplete, abstract code that makes predictions about the world and revises its predictions on the basis of interactions with the world" (Lauwereyns 2012: 74).

To me, the overall sense conveyed by this passage has a lot to do with the broad theoretical context of AORs. Indeed, saying that PP's representations "are not actual representations" while preserving the concept of representation seems to be a clear endorsement of what Clark previously called "minimal robust representationalism (MRR)" (Clark, Grush 1999: 6-8), of which action-oriented representation is the core concept.

In the 1999 article, Clark and Grush talk about MRR to explain how real-world, real-time actions, like skilled reaching for example, are performed successfully.

The two authors start to consider this problem saying that the success of smooth, world engaging actions depends on the brain's ability to respond to proprioceptive feedbacks (concerning the position and trajectory of the hand system, for example). After having noticed that often those feedbacks are

required faster than they are available (they are required before the minimum delay has elapsed), and having suggested that the system needs something more than those feedbacks to engage with the real-time action quickly, Clark and Grush draw a comparison between the human proprioceptive system and emulator mechanisms to explain this phenomenon.

Emulators are mechanisms that

i) take information about the current state of a system and about control command issued as input,
ii) and give a *prediction* of the next state of the system as output.

This mechanism is said to be crucial for the success of a system's task, because it "allows the system to mock feedback signals available ahead of the real-world feedback, and hence allows rapid-error correction and control" (Clark, Grush 1999: 6).

Anticipatory mechanisms described in Clark and Grush's 1999 article are said to rely on a particular form of representationalism: minimal but robust representationalism. This form of representationalism, similarly to what noticed in the quotation from Clark's "Radical Predictive Processing" (Clark 2015c), is said to be minimal because, in contrast to full-blooded representations, representations computed by the system are not fully decoupleable, and do not offer a rich description of the external world (Clark, Grush 1999: 6-7). Nevertheless, those representations are said to be robust because they are internal states that do a great job in guiding action, are scientifically identifiable, and serve as stand-ins for specific extra-neural states of affairs (Clark, Grush 1999: 10).

As previously said, those representations are said to play a central role in guiding action because

i) by mocking feedback signals available ahead of the real-world feedback, they guarantee the action-perception based cognitive performance to be smooth and quick;
ii) they allow rapid error control.

This description representations in emulator-like systems is theoretically really close to anticipatory mechanisms Clark considers in his PP. Indeed, like what happens in the description of emulator mechanisms, PP it said to guarantee a smooth engagement with the world because the brain's internal

system represents anticipations of feedback signals, and it widely relies on mechanisms of error minimization (PEM).

To me, the crucial point here is the way Clark extends the seminal idea of error control described in the 1999 article in his work on PP.

As explained in the previous paragraph, by means of the concept of active inference, Clark strongly emphasizes how action contributes to PEM mechanisms. Nonetheless, despite this strong emphasis on embodied action, the body and its possibilities of action are taken into account in an indirect way. Indeed, as pointed out in §III.4, action is considered to be crucial in the PP framework because it delivers fragments of information that the brain uses to control subsequent action (Clark 2016: 250).

My philosophical point consists in claiming that the way Clarks thinks of these fragments of information delivered to the brain by action, and the way this information fits in generative models is the same suggested by MRR by appealing to the concept of AOR. This is to say, AORs Clark talks about in his PP are the same kind of representation he talked about in 1994 (Clark, Toribio 1994), 1997 (Clark 1997a), and 1999 (Clark, Gursh 1999).

At a general level, this is confirmed by Clark's definition of the brain's generative models as "context-sensitive generative models that simultaneously prescribe recognition and action" (Clark 2016: 125). This sounds like a clear reassessment of Clark's first approach to AORs, offered in his 1997 book. Indeed, there (Clark 1997a: 49), recalling Millikan's idea of Pushmi-Pullyu representations, he said that AORs are neural populations that simultaneously describe aspects of the world (= recognition aspect) and prescribe possible actions.

Moreover, he described AORs as "action-dependent", and as reflecting "the profound role of bodily motion [...] in shaping and simplifying the information-processing problems to be solved" (Clark 1997a: 149).

Also this second aspect, namely AORs' action-dependence, reflection of bodily motions, and simplification information is consistent with Clark's broad idea of generative models in relation to active inference. Indeed, he states that PP is action-oriented (Clark 2016: 111), he accounts for bodily motions in internal models by focusing on the epistemic role of feedback signals conveyed by the senses while agents perform world-engaging tasks (Clark 2016: 190, 247, 256), and he says that action is important because it makes the brain's informational load lighter, fostering predictive mechanisms and PEM.

Furthermore, Clark himself -even if sporadically- makes use of the phrase "Action-Oriented Representations", and he does so when he deals with active inference. For example[121], he says that

> "many of the probabilistic *representations* inhering the generative model will now be [...] '*action-oriented*'. They will represent how things are in a way that, once suitably modulated by the precision weighting of prediction-error, also prescribe (in virtue of the flows of sensation they predict) how to act and respond. They are representations of affordances" (Clark 2016: 133; italics added)

Later, he defines those representations as "pragmatic representations", borrowing the expression from Cisek and Kalaska's article "Neural mechanisms for interacting with a world full of action choices" (Cisek, Kalaska 2011). The way Clark describes these pragmatic representations is the following one:

> "representations tailored to the production of good-online control rather than aiming for rich mirroring of an action-independent world. Those representations simultaneously serve epistemic functions, sampling the world in ways designed to test our hypotheses and to yield better information for the control of action itself" (Clark 2016: 251)

Taken together, those passages suggest a well-known philosophical story about action: the epistemic role of action (and the control of action too) should be individuated in the way *possibilities of action* are *represented by the brain*.

As explained in the previous paragraph, Clark says that action contributes to PEM because it produces samples of the world. The second passage from *Surfing Uncertainty* cited above makes this story about action more complete, individuating the nature and the location of those samples of the world. Indeed, pragmatic representations embedded in the brain's generative models are the result of the job of sampling active inference is said to do. Then, it is not action in itself that is central in anticipatory processes that control the agent's engagements with the world, but again action-oriented representations (samples) provided by action.

To Clark, the fact that those representations are pragmatic (or action-oriented) would allow his explanation to close the distance between his PP and enactivist approaches to cognition (Clark 2015c: 20), and, crucially, this would

[121] See also Clark 2015c: 21.

put an end to representation wars (Clark 2015a). Given that representations he makes use of in his PP have the same nature of representations he described in his previous work on action-perception loops, and considered the objections against this kind of representation developed in the "enactivist literature", the "peace treaty" Clark suggests enactivists to sign seems not being acceptable.

This not only because Clark's PP does not account for action as a "doing", and this is what enactivism thinks we should focus on (Gallagher 2017: 19), but also because the epistemic posits Clark relies on to explain the role of action in predictive processes has been shown to be philosophically unwarranted. As AORs of *Being There*, pragmatic representations of Clark's PP do not fulfill the representational criteria usually endorsed to talk about representations properly. In fact,

i) neural populations do not fulfill the standing-for and the interpretability condition;
ii) pragmatic representations are not decoulpleable from what they represent, because they are said to occur in on-line control (Clark 2016: 251).

Therefore, the use of a representational vocabulary in Clark's PP, considering "probabilistic representations", "error representations" and "pragmatic representation", is unwarranted (Bruineberg, Kiverstein, Rietveld 2016; Gallagher 2017; Orlandi 2016). This not only makes Clark's project of moving away from the classical representational paradigm (Clark 2015a; 2015c; 2016) unsuccessful, but it also undermines his claim about the possibility an "extended" account of predictive mechanisms (Clark 2016; 2017).

In fact, the appeal to internal representations, in particular to AORs or pragmatic representations, is expressive of an internalistic prejudice, namely the idea that embodied action matters in predictive processing because it is encoded in the brain's generative models. This prejudice -which is also hidden in the implicit appeal to the Parity Principle- prevents Clark's explanation from *actively extending* predictive processing in the real-world. Clark's explanation, like in his minimal robust representational proposal developed in the 90s, underplays the epistemic power of real-world embodied action (namely world-engaging cognitive practices). If one the one hand it is true that both Clark's ways to conceive of action-oriented representations (that of the 90's and that of PP) are attempts to describe what representations *do* in cognitive processes (Ramsey 2007: 18; i.e. prescribing and controlling action), and this seems to

point to an active account of representations, on the one hand it is also true that this way to conceive of action is weak. In fact, action prescription and action control, as claimed in §III.4, §II.5, and §II.7, account for conative aspects of cognitive experience and not for action itself.

III.6 The enactive approach to Predictive Processing: an alternative proposal.

In the last two paragraphs, I considered Clark's approach to Predictive Processing, examining the main points of his proposal, aimed at reassessing predictive mechanisms in order to make them meeting his EMH. In particular, I noticed that the way he tries to extend PP relies on some core concepts: multiplicity and flexibility of Markov blankets, Parity Principle, and centrality of action conceived in terms of active inference. My discussion has shown the interdependence of those concepts, which are grounded on an informational account of cognitive processes. In Clark's proposal, this informational conception of cognition is said being able to extend PP because it allows the explanation to individuate Markov blankets' boundaries where there is information processing, independently to the internal location of information. Moreover, it allows action to be a crucial part of anticipatory cognitive mechanisms because action is said to provide the cognitive system with fundamental information for prediction error minimization.

I suggested that, even if at first glance this way to think of PP sounds like an appealing strategy to extend this explanatory framework the explanation does not actually succeed in its project of extending the predictive mind. This not only because Clark still sticks to a philosophical principle that hides an internalistic prejudice (Parity Principle), but also because a careful exam of Clark's PP at work shows the centrality of the concept of representation in his explanation.

An analysis of concepts such as "error representation" and "probabilistic representation" shows that the use of these concepts is unwarranted, because the mechanisms Clark talks about do not fulfill the necessary conditions to identify representations properly. This is philosophically problematic, but what is even more problematic is the use of the concept of "action-oriented representation". Indeed, not only those items or processes are not genuinely representational (because of the same reasons why "error representations" and "probabilistic representations" are not representations), and then the use of a representational vocabulary to describe them is not justified, but the very idea of "action-oriented representation" is particularly dangerous in the PP framework. In fact, Clark claims that PP is suitable to combine the EMH with the enactive approach to cognition because it explains anticipatory mechanisms by widely focusing on action. Given that in Clark's PP action is explained through the concept of AOR, this philosophical project seems unsuccessful. In fact, as claimed previously, AORs i) do not actually account for action in itself,

and ii) they "crystallize" action in internal cognitive mechanisms, to which an epistemic primacy is implicitly given. Therefore, instead of extending the predictive mind -claim that entails that the mind extends beyond the brain by means of embodied action- this explanation explains the mind just as enhanced or scaffolded by external action: this suggests that cognitive mechanisms have an internal core, *plus* external parts.

To extend the predictive mind, a crucial role to the very cognitive processes that take place outside the head should be given. Moreover a stronger emphasis on the very relation between internal and external dynamics is required. To do that, since the notion of representation, and in particular the concept of AOR, has been shown to entail internalistic consequences, the explanation should not give AORs "new vehicles to ride" (Hutto 2013: 145). On the contrary, it should get rid of them, and focus on embodied action in a stronger sense. The enactive account of Predictive Processing[122] precisely aims to tell this story.

In this paragraph, I will discuss this approach to Predictive Processing (sometimes called "Predictive Engagement"; Gallagher, Allen 2016; Gallagher 2017). This discussion aims to offer an alternative story of cognitive anticipatory processes that puts an end to representation wars (Clark 2015a), avoiding this epistemic posit when it is not necessary, namely in the explanation of action-perception loops, in which anticipatory mechanisms play a key role.

The main philosophical point at the core of the enactive explanation of PP concerns the place that the brain occupies in anticipatory dynamics (Gallagher, Bower 2014; Gallagher, Allen 2016; Gallagher 2017). As explained in §III.3 and §III.5 -to different extents- both Hohwy's Predictive Coding and Clark's Predictive Processing account for anticipatory mechanisms of cognitive processes starting from an analysis of the brain's functions. Both the explanations conceive of the brain as a pro-active machine that, anticipating its own sensory states thanks to the job of internal models, and engaging in operations of prediction-error minimization, allows the cognitive agent to act in her environment smoothly and quickly.

The story about the brain enactive approaches to PP draw is similar *but* different. As briefly noticed, also enactivism emphasizes the role of the brain in anticipatory processes. Nonetheless, contrarily to what happens in the two

[122] Notice that there is not a unitary enactive account of PP, even if enactivist scholars focus on the same theoretical points (e.g. the centrality of action in prediction, the non-representational nature of PP). Here I sum-up the core ideas suggested in some enactivist papers (Bruineberg, Kiverstein, Rietveld 2016; 2014; Gallagher, Bower 2014; Gallagher, Allen 2016; Hutto 2017) in order to illustrate the broad enactive account of PP.

explanations already considered, it does not give an epistemic primacy to those internal mechanisms. This because the epistemic primacy ascribed to the brain hides an internalistic and conservative prejudice, and it usually leads to conceive of predictive processes as representational. In fact, if the brain is thought to be the central part of these mechanisms, it usually follows that

a) the overall description of anticipatory processes implicitly or explicitly entails that cognitive success depends on internal models, where the idea of internal model (also when models are said to be "frugal" and "action-oriented") suggests a demarcation between the brain and the environment that sounds not suitable to account for the *situatedness* of anticipatory mechanisms;
b) internal anticipatory and PEM mechanisms should make use of representations of some sort, because the brain, in order to "communicate" with the body and the world, has to process information conveyed by the senses by working on the contents of these fragments of information, and it has to give motor commands to the body, by representing what should be done upon aspects or features of the world.

A different way to think of the role of the brain in anticipatory mechanisms is offered by Gallagher and Bower (Gallagher, Bower 2014) and by Gallagher and Allen (Gallagher, Allen 2016). By generalizing the "Interactive Brain Hypothesis" developed by Di Paolo and De Jaegher (Di Paolo and De Jaegher 2012)- which has been formulated to explain the brain's mechanisms for social understanding- the enactive approach conceives of the brain just as *one* of the components of a process of dynamic attunement (Gallagher, Bower 2014: 233). This kind of dynamics involves the brain, the body, the physical, social, and cultural environment. As the phrase "Dynamic attunement theory of brain" suggests, this process is aimed at settling each component of the system in the right kind of attunement with the others. This right attunement guarantees the brain-body-environment system to be in a state of equilibrium that, if considered at the cognitive level, allows the agent to acquire a grip on her environment.

Considering this point at an evolutionary level, it can be said that the brain can be considered to be just one of the components of the larger system. This not only because the environment and the body do a great job in cognitive processes, and function as its enabling conditions (i.e. they determine the functionality of the cognitive performance; Di Paolo, De Jeaegher 2012: 5), but

also because the brains we have do what they do and function in the way they function because they have been shaped during evolution by the bodies we have. The fact that humans have hands ready to grasp, and eyes structured to focus determines the way our brains evolved, and this shapes the way they work in tandem with our bodies (Gallagher, Bower 2014: 240).

According to the enactive perspective, this is the crucial starting point a theory of PP should be grounded on in order to overcome the problems the accounts previously examined entail. Even if at this point of the explanation it can seem that saying that the brain is part of a larger system is not enough to really account for the role of the brain in anticipatory processes, I think that this suggestion is crucial to consider one of the main problems of the PP framework: the appeal to internal representations embedded to the brain. Indeed, the way this perspective on the organism rethinks of the fundamental explanatory unit of the theory (the whole organism-environment system; Gallagher, Allen 2016) lays the foundations of a non-representational account of the brain. If the brain is thought to be just one of the components that work in tandem in the same system, the explanation can avoid to recall the notion of representation, and taking *circular causality* as the *explanans* of the process (Gallagher 2017: 129; Bruineberg, Kiverstein, Rietveld 2016). To put in other words, an explanation that gives an epistemic primacy to the brain, and considers it as an individual system implicitly entails that the brain is ruled by a different kind[123] of law than that of the body and the world. It follows compositional and interpretative rules that allow it to represent what is outside of its system, namely the world and the body, and then to have a cognitive grip on the other systems. Conceiving of the brain as part of the same system of the body and the environment allows the explanation to appeal to the same rules to explain how different components of the larger system function. In the case I am taking into account, this means explaining what the brain does saying that its activity is affected by changes produced in the environment and the body, which are part of the same system, guided by the same law (causality). On its turn, changes in the brain affect the body, and the agent's possibilities of action, which allow the agent to achieve a grip on her world. To put in other words, conceiving of the explanatory unit of the theory as an integrated system made

[123] This does not entail that explanations that give a pivotal role to the brain in cognitive processes deny a causal explanation of the brain's mechanisms. Nevertheless, when they get to explain the cognitive contact of the brain with the world, they are drawn to appeal to representations, because they have to account for the mediation between two systems different in their nature: the body-environment, and the neutral system. For this point, see also the discussion about the problem of representation developed in §I.7.

of the brain, the body, and the environment, where no epistemic primacy is given to one of the components, means thinking of the components of the system as undergoing processes that co-vary the one with others (Bruineberg, Rietveld 2014: 7)[124]. Given this relation of reciprocal causality that links the parts of the fully embodied system, it no longer makes sense to postulate that the brain needs to infer the causes of its owns stimuli, or that it works on internal samples of the world in a way that is suitable to test its own hypothesis. Rather, anticipatory processes are explained as the co-variance, adjustment, or *attunement* (Gallagher, Allen 2016; Gallagher 2017: 163) of the parts of the larger system. The brain, as the other parts of the system, just attunes itself to the other components, and, on its turn, allows the body and the environment to be attuned the one with the other[125].

What the "Dynamical attunement theory of the brain" entails -namely the idea that the brain is part of a larger system- can be reassessed by considering the way the enactive approach to predictive mechanisms interprets one of the key concepts of the Predictive Coding theory: *model*.

As explained in the previous paragraphs, Predictive Coding in Hohwy's version, and in Clark's one as well, widely relies on the concept of internal model. Howhy's proposal emphasizes the ability of the brain to embed the probabilistic structure of the causes of its own states, and Clark's approach focuses on the action-orientedness of these models, which simultaneously serve a function of recognition and a prescriptive function. What those approaches share is the point where the boundaries of those models are individuated: in different cortical areas.

The enactive approach to PP suggests that, in order to avoid the internalistic consequences others explanations of anticipatory mechanisms entail, we should think of the boundaries of those models in a different way. In the same way the "Dynamic attunement theory of the brain" suggests that the explanatory unit of the theory should be individuated in the brain-body-environment

[124] Notice that the notion of covariance is also discussed by Hutto (Hutto 2017) in order to give a non-representational account of information processing, which allows the explanation to avoid a description of the brain as processing representational contents. Indeed, covariance explains the job of the sensory system as being set up to respond to aspects of the world that track specific kinds of states.

[125] Such a dynamics can be also explained by making use of the concept of "intrinsic activity of the brain". The idea is that what the brain requires at any given moment depends on its self-generated activity, which is supported by the whole activity of the life-sustaining system (Thompson, Cosmelli 2011: 169-170; 172).

system, the enactive approach to PP claims that the model PP relies on is constituted by the organism itself, which embodies "the longer-term regularities between action, environment and the state of the organism" (Bruineberg, Rietveld, Kiverstein 2016) in its own structure. This means that, instead of individuating different generative models at the levels of the hierarchical structure of the brain, the model PP makes use of a larger model, which encompasses the brain and the body. Moreover, this model does not need to represent what lies outside of it. Indeed the organism is not said to *have* a model of the external world, of its causal structure, or regularities. But, on the contrary, it is said to *be* the model of the way those regularities are connected to experiential regularities: the organism is a model of its Umwelt because it enacts it. When it acts upon its environment, the organism enacts regularities of the subjective world, and it does so according to its possibilities of action determined by skills, which model what can and should be done in the world in order to preserve the life-continuity of the organism.

This broader interpretation of the concept of model not only recalls the idea of coupled system (agent + Umwelt) offered by Varela, Thompson and Rosch in the milestone book *The Embodied Mind*, but it is also consistent with one of Friston's ideas. Indeed, Friston himself says that "an agent does not have a model of its world, it *is* a model" (Friston 2013; Bruineberg, Kiverstein, Rietveld 2016). Therefore, he acknowledges the possibility of making use of the concept of model in a broad sense, thinking of the fully embodied agent as modeling its econiche, namely as embodying its subjective world, the organism itself is constitutive of.

It seems to me that this represents a great shift from Hohwy's conception of model, and from Clark's one too. As a matter of fact, even if Clark, by postulating multiple and flexible Markov blankets that encompass the boundaries of the skull, extends the components of the PP system, he does not endorse the full claim according to which models PP makes use of are neural and extra-neural, namely that models *coincide* with the whole organism. Also when he takes into account the problem of sampling, which is a product of embodied action conceived as active inference, Clark thinks of samples of the world as parts of action-oriented anticipatory and probabilistic models, which are said to be located in the brain.

Clark's narrow consideration of the concept of model seems to depend on the lack of attention he pays to the philosophical implications of the notion of free-energy, and on the underestimation or on a misleading interpretation of concepts related to that of free-energy, above all that of action.

The enactive account, by emphasizing the philosophical implications of the concept of free-energy, is able to endorse the premises of a more extensive account of PP.

The enactive approach to PP (Bruineberg, Rietveld 2016; Bruineberg, Rietveld 2014) begins its discussion about free-energy noticing that this principle -which other theories of Predictive Coding interpreted in relation to the brain's dynamics, saying that the brain always tries to maintain itself in limited sets of states in order to minimize the dispersion of free-energy- can be applied to the whole dynamics of the organism. To me, this makes a lot of sense, considering the cognitive phenomenon Predictive Processing aims to explain. In fact, this theory wants to explain how agents acquire a grip on their worlds in on-line cognitive practices (as Clark clearly points out; Clark 2016). On-line cognitive practices are situated in a context in which the agent, thanks to her body and skills, does something upon features of the world. Then, the state of equilibrium the theory should account for is not the state of equilibrium of internal dynamics alone, but the equilibrium between internal *and* external dynamics.

According to the enactivist approach to PP, free-energy is exactly the measure of the disattunement between internal and external dynamics: it measures energy dispersion in exchanges between the activity of the brain and the body, which, together, constitute and integrated system.

Crucially, the concept of free-energy is strictly tied to the embodiment of the system. Indeed, for a given system, being in a state of equilibrium depends on the survival needs of the organism. Those needs are defined by the way the living body determines the "best" situation of the organism in the environment.

For example (Bruineberg, Kiverstein, Rietveld 2016), think of a whale in deep sea. In this case, the whale finds itself in a state of equilibrium, or low suprisal, because it is situated in the habitat where it is supposed to be, namely in a habitat suitable for the preservation of its life and for the performance of its life-sustaining activities. If the organism in deep sea is a human, the situation is different. The human organism's conditions of equilibrium are different, then in that situation, the degree of free energy would be higher; the organism would find itself in a state of high surprisal. This not because the human's internal models give a wrong internal representation of the causal structure of the world, but because the broad model (that is the whole organism) is not attuned to the present environmental conditions.

This disattunement can be formalized in terms of surprisal, as the classical version of Predictive Coding does, namely in terms of probability distributions.

Nevertheless, as the enactivist proposal I am considering suggests, there is a another way to think of this disattunement, which can be helpful to consider in the economy a theory that seeks to account for what an organism actually does in order to keep itself in a state of equilibrium (and this means minimizing free-energy). The concept this approach to PP appeals to is that of *embodied surprisal*. As previously explained, in the Predictive Coding theory, suprisal is different from conscious surprise. Suprisal is not experienced, but it is rather a state of the brain, which signals and measures mismatches between predictions and incoming signals. The point is that, if we have to consider the overall predictive dynamics of the organism, the subpersonal conception of surprisal seems to narrow. This not because it is not explicative at all, but because it does not account for the very connection between a neural phenomenon and a cognitive and personal-level phenomenon, namely the activation of strategies aimed at reducing the disattunement between internal and external dynamics. It is not clear how the embodied, experiencing agent can have access to fragments of information about a neural state of disequilibrium, knowledge that would dispose her to display strategies (i.e. action) for the minimization of disequilibrium. Therefore, the enactivist proposal suggests to integrate this crucial notion of surprisal with "embodied surprisal", namely with what the agent consciously experiences when her predictions are not fulfilled by what happens in the on-online engagement with the world.

This mismatch between predictions and sensory experience is interpreted by the enactive explanation in terms of *affective tension*. The idea that lies at the core of the concept of affective tension is theoretically connected to the authors' previous work on perception and affordances (Kiverstein, Rietveld 2012; Rietveld 2008a; 2012a; Rietveld, Kiverstein 2014), discussed in Chapter II. There I explained that these enactivst scholars conceive of perceptual experience and action as always and fundamentally affectively caractherized, and this depends on the crucial concept of an afefctive-pragmatic intentionality, which is said to pre-reflectively guide the agent's engagement with her environment.

This approach to action and perception relates to this enactivist work on PP because the generative models PP is said to rely on are through to be a nest of *states of action-readiness*. These patterns of action readiness are the agent's internal states that, according to her sensory states and abilities (skills), prepare her to achieve a grip on the environmental situation. Action-readiness - according to this perspective that integrates a neuroscientific explanation within a phenomenological framework- is *affectively connoted*. In fact, even if this

explanation does not say that action and perception are always forerun by goals, it accounts for the directedness of action and perception in terms of needs and concerns, namely in terms of what the agent cares about.

This is to say that the agent is ready to do something upon an aspect or feature of the environment because she cares about those aspects of the world, which perceptually appear as attractive and salient. This affective phenomenon of caring about the world, in Chapter II, has been described by referring to the needs, projects, or interests that move the organism in the changing situation of its environment (Rietveld 2008a: 342). I suggest that this phenomenon can be explained in terms of what Colombetti calls "*primordial affectivity*" (Colombetti 2014: 11), reassessing Heidegger's concept of *Sorge* (Heidegger 1927). Primordial affectivity, rather than being a conscious emotion, is a property of a

> "specific organization that sets up an asymmetry between the living system and the rest of the world, which consists in a perspective or point of view from which the world acquires meaning" (Colombetti 2014: 2)

Primordial affectivity is the basic affective tone the agent's encounter with the world is structured by. It structures experiences and dispositions to act according the point of view of meaning, where meaning is basically understood in terms of "life values" (e.g. well-being) of the organism.

This suggests that there is a tension, a "directedness" toward a certain, *expected and predicted grip* on the environment, aimed at producing the organism's well-being. This grip on the environment is expected in the sense that it is *affectively anticipated* by the perceptual system, which is open to or attracted by certain saliences of the environment in virtue of this basic primordial affectivity or care (Rietveld 2008a). Moreover, the agent's grip is also said to be predicted by the brain itself in an affective way (Rietveld, Kiverstein 2014). Indeed, as anticipated in the previous chapter, recent research (Holroyd, Coles 2002; Barrett, Bar 2009[126]) show that the brain, thanks

[126] For a more detailed explanation (and for a non-representational interpretation) of Barrett and Bar's research see Gallagher 2017: 19 and ff., 166 and ff. Notice that also Clark, in his latest work on PP, considers Barrett and Bar's "affective prediction hypothesis", but he draws conclusions that differ from the enactive approach to the role of affects in PP. Indeed, even if Clark acknowledges the crucial role of the affective dimension in anticipatory cognitive mechanisms, he does not question the representational status of affect (Clark 2016: 165), which Barrett and Bar seems to take for granted when they talk about "affective representations" of the brain (Barrett, Bar 2009: 1325). As it will be discussed later, the way the enactive approach to PP describes affects is non-representational, and it is phenomenologically characterized.

to its ability to recall the affective impact of sensations of the past, makes predictions about the meaning of visual sensations in the present. This allows it to "anticipate and prepare to act on those sensations in the future" (Barrett, Bar 2009: 1326).

When states of action readiness do not allow the agent to achieve the expected grip on her environment, namely when the agent's states of action-readiness do not encounter environmental conditions suitable for the performance of the expected action, the agent experiences an affective tension: she feels that something has to be improved.

Even if Bruineberg, Rietveld and Kiverstein (2016) and Rietveld and Kiverstein (2014) do not explain affective tension exactly in this way, it can be helpful to flesh out their argument by considering one important concept in the affective sciences: *appraisal*, namely the "process that detects and assesses the significance of the environment for well-being" (Moors, Ellsworth, Scherer, Frijda 2013: 120).

According to standard cognitive theories of emotions (e.g. Lazarus 1991; Smith, Lazarus 1993), emotions contain evaluative judgments (appraisals) about their formal objects (Prinz 2003: 54). Formal objects of emotions are sometimes defined as "core relational themes". Those ones -as Lazarus explains in his *Emotion and Adaptation*- are recurring topics that express "ongoing relationships between persons and their environments" (Lazarus 1991: 22); they are "things that are of value to the organism" (Hufendiek 2017: 4459). Such relationships are expressive of the personal harms and benefits of the relation between the agent and the environment (Lazarus 1991: 39). To appraise core relational themes means -for example- explaining fear as the judgement "there is danger", then fear is thought to be dependent on the harm that characterizes the relationship between the agent and her environment.

Appraisal of emotions, because it is an assessment of the relationship between the feeling agent and her environment, has an adaptive function. Since the agent knows something about her own emotions, she is prepared to cope with the situation of harm, for example, minimizing or alleviating the appraised harm. Crucially, the appraisal of an emotional circumstance depends on the agent's goals and beliefs. In this sense, appraisal mediates between internal states of the agent, and states of the environment, preparing the agent to change her environment in order to maximize the benefits of her relation with it (Smith, Lazarus 1993: 234).

As Colombetti points out, standard theories claim that emotion appraisal takes place in the agent's head (Colombetti 2007). It is considered to be a

process that operates on a wide range of representations (conceptual, propositional, perceptual, embodied, and so on), whose contents are appraised values (Moors, Ellsworth, Scherer, Frijda 2013: 120-121; Prinz 2004, Chapter 3). Nevertheless, some contemporary views on emotion appraisal affiliated to the enactivist lineage (Colombetti 2014; Colombetti 2007) challenge this internalistic and representational conception of appraisal. Instead of considering appraisal processes as taking place in the head, and as just interacting with the body, conceived as a means of appraisal thanks to its capacity of being emotionally aroused, the enactive theory of emotions conceives of the body as a *constitutive* part of the appraisal process. According to this explanation, the activity or behavior of the organism is an activity of appraisal by itself (Colombetti 2014: 101), and the aim of this activity is to maintain the equilibrium between the agent and its Umwelt.

According to Colombetti (Colombetti 2007), emotion appraisal is constituted of many factors: the physical systems that underpin perception, physical processes that take place in the brain, proprioceptive and kinetic experiences, felt arousal, and action orientation. This suggests that the appraisal of emotions -which, considering what Lazarus says, pertains to basic relations between the agent and the environment (e.g. harm or benefit)- takes place in a larger system. The activity of this system is distributed between a body aroused by emotions, a body kinetically connoted, and that does something in the environment according to this "corporeal feeling", and the environment itself, which displays affective features.

This means that, in this perspective, "core relational themes", which affectively inform the agent about her relation with the environment, do not acquire a cognitive status in virtue of their being represented or judged by the agent. On the contrary, basic emotions that trigger the agent's behavior are *enacted*, namely, they are brought forth by the encounter between the world and the embodied organism (Colombetti 2014: 110-111).

This embodied and enactive description of appraisal offered by Colombetti, is helpful to make clear what Bruineberg, Kiverstein and Rietveld (2016) and Rietveld, Kiverstein (2014) mean what they describe embodied surprisal as an "affective tension" or as "the feeling that something has to be improved". Indeed, the phenomenon they talk about concerns something that is very similar to what the notion of "appraisal of core relational themes" expresses: the basic affective situation of which the constituents are an embodied organism and the environment. Moreover, the consequences the notion of "affective tension" entails and the appraisal of core relational themes, in particular if

appraisal is conceived in the way suggested by Colombetti, seem to be the same. The embodied appraisal of the affective tension the agent feels draws her to act in a certain way in the environment, in order to minimize the situation of disequilibrium between internal and external dynamics, affective tension conceived as the appraisal of a core relational theme is expressive of. To put in other words, the affective tension is an enacted response (appraisal) to core relational themes (Hutto 2012: 179; Colombetti 2014: 110-111), and this response entails: i) bodily feelings and changes; ii) dispositions to act, which will form part of coping-strategies.

What is particularly interesting in thus enactive approach to PP is the way the explanation accounts for the consequences of the embodied and situated appraisal of affective tension.

The authors (Bruineberg, Rietveld 2014), by recalling the notion of "field of affordances" (Rietveld, Kiverstein 2014), give a lot of theoretical space to the role of the "affectively tuned" econiche to explain how action, crucial for predictive error minimization, takes place. They claim that what draws the agent to improve her grip upon the context of the cognitive practice -namely to minimize the disequilibrium between her states of action-readiness and what she can actually do in the environment- is not a reassessment of the hypotheses represented by the brain's models (Hohwy's proposal), or, like in Clark's account, an embodied action guided by internal action-oriented representations that "say" what to do and how to do it. Rather, the world itself -conceived as a rich field of relevant affordances- fulfills this epistemic function (Bruineberg, Kiverstein, Rietveld 2016; Bruineberg, Rietveld 2014). To minimize embodied surprisal – which is expressive of the disattunement between states of action-readiness and contextual possibilities of action– the agent is solicited by environmental affordances to perform actions that will provide the sensory system with the right kinds of perceptual states, and this means making internal states of action-readiness attuned with on-line contextual action.

At first glance, this point of the enactivist proposal sounds similar to Clark's account of action, because this philosophical proposal seems to suggest that action minimizes embodied surprisal. Nevertheless, a closer look at this proposal shows that there is a very important difference between these two accounts, and this difference -I claim- makes the enactive proposal more interesting than Clark's one. Indeed, in Clark's account, since active inference is said to rely on AORs, the direction of the process aimed at establishing the equilibrium between internal and external dynamics goes from the internal

mind towards the world[127], and only later from the world towards the mind. As already claimed, this undermines Clark's proposal of an extended predictive mind, because the concept of AOR entails that the role of real, embodied and *contextual* action is underplayed.

The enactivist explanation, thanks to the emphasis on the idea of the field of relevant affordances, avoids this problem. Indeed, this approach to PP claims that "we can learn something about the brain [namely its anticipatory mechanisms] by investigating the structure of the econiche" the organism acts upon (Bruineberg, Rietveld 2014: 9, text into brackets added). The very structure of the landscape of affordances triggers patterns of action that shape the agent's action-readiness, namely the organism's anticipatory mechanisms. Crucially, the field of relevant affordances shapes anticipatory mechanisms in an affective way. Indeed, as explained in Chapter II, econiches not only offer the agent the perceptual substrate that affords skillful action, in the sense that they offer possibilities of action perceptually conveyed, but they provide affective features crucial for action-anticipation and action-selection.

Even if the enactivist work I am considering does flesh out this point by connecting affordances, affect, and core relational themes, I think that if this this theoretical connection is taken into account, the affordance-based argument about PP can be understood better.

As explained in the discussion of the enactive interpretation of the concept of appraisal, when the enactive affective science conceives of core relational themes tends to "externalize" them (Hufediek 2017: 4456; Hufendiek 2015: 106). By emphasizing the adaptive function of core relational themes, it conceives of them not just as felt emotions or as representations of occurrent emotions, but rather as the enactment (appraisal) of a basic emotion in virtue of an embodied action performed in the environment (Colombetti 2014: 110). In this sense, core relational themes do not seem to pertain to the agent's private emotional experience, but rather the relation between an affectively tuned environment (e.g. an environment expressive of loss, harm, danger, and so on[128]) and the agent's response to that environment. This leads to think of basic emotions (such as the feeling that something has to be improved) as *situated* in a cognitive practice, where the cognitive and practical nature of

[127] Notice that Bruineberg, Kiverstein and Rietveld (2016) raise this objection against Hohwy's Predictive Coding. Nevertheless, considered what said about internal action-oriented predictive models, which are said to prescribe action, I think that this critique applies also to Clark's proposal.

[128] For more examples of "core relational themes", see Lazarus 1991: 122, table 3.4.

emotional experience depends on the adaptive function of emotions and on their ability to trigger some kind of action or behavior.

This way to conceive of core relational themes, or of "affective tension", as situated, is particularly interesting to understand Bruineberg and Rietveld's affordance-based explanation (Bruineberg, Rietveld 2014). Indeed, if we conceive of core relational themes as situated, we can integrate one of the authors' core insights: action is guided by features of the environment that afford an emotional response, crucial for action selection, action guidance, and the anticipation of action(see also Rietveld 2008a). Core relational themes, in virtue of their situated nature, support this affordance-based explanation because: since core relational themes are situated patterns of reactions, they sound like affordances affectively connoted, or "affective affordances".

This is crucial for the enactive explanation, because it accounts for preparatory and anticipatory aspects of action from a non-internalist perspective. The enactive explanation claims that action, which in the PP framework is crucial for PEM, is prepared by affects, which are triggered by relevant affordances, namely affordances that stand out in the field of affordances because they are "affectively attractive". Conceiving of core relational themes -namely the affective tension the authors talk about when they discuss embodied surprisal- as situated explains the affective relation between anticipation, emotion, and rich the field of affordances better. The claim according to which an optimal grip on the field of affordances, which allows the agent to act as "surprisallessly" as possible (Bruineberg, Rietveld 2014: 10), is guaranteed by anticipatory mechanisms that have to do with affects and emotions (Bruineberg, Rietveld 2014: 10; 11) can be fleshed out in this way. The landscape of affordances triggers emotional responses that shape the agent's action-readiness patterns (namely anticipatory mechanisms) because this landscape is structured around core relational themes, namely recurrent relations between the agent and the environment expressive of basic emotions connected to the "good" or "ill" of an organism-dependent situation. This is to say that the anticipatory strategies followed by the agent in her cognitive practices are shaped by an affective landscape of affordances because

i) the whole organism, thanks to evolution, "is set up to be set off by core relational themes by responding to these" ones (Hutto 2012: 179).
ii) Prediction error minimization -namely strategies suitable to attune internal states of action-readiness with external dynamics- is actively performed in the rich landscape of affordances in virtue of an

affective anticipation of action, core relational themes are expressive of.

iii) Prediction error minimization is supported by affective dynamics. Those ones characterize the very relation between an embodied agent and her environment in virtue of the affective structure of the landscape of affordances, in which some affordances stand out more than others because they are affectively attractive (namely they offer possibilities of action for the organism's good). The attractiveness of those features of the environment drives the agent to act on these selected environmental features, which will fulfill the agent's predictions, namely action-readiness patterns.

What is crucial in this explanation –and this is what actually makes the difference between the enactive account of PP and Clark's one in their project of extending PP– is the way affect and action are nested in anticipatory mechanisms.

In Clark's proposal, embodied action and externalities come into the explanation with the notions of representation, internal processing of information, and internal model. This makes the explanation closer to more classical models of Predictive Coding, because an epistemic primacy is given to internal cognitive resources and mechanisms.

In the enactive account of PP, the controversial notion of internal representation (in particular that of AOR), which in Clark's work is said to explain the anticipation of action, is replaced by the concept of affect. What is crucial is that affect is not thought to be a property of the brain (even if affect has something to do with neural dynamics, in particular in the "affective prediction hypothesis") or a personal feeling-quality of the agent. Rather, it is considered to be "situated" in the relation between the whole organism and the environment, where this relation, in its being adaptive, is intrinsically active, since action is establishes the equilibrium between internal and external dynamics. Therefore, this approach to PP gives a more extensive account of the "predictive mind" than Clark's one because it puts at the core of its explanatory framework a concept (affect) that, appropriately understood, explains the anticipatory underpinnings of action-perception loops as embodied and situated.

III.7 Concluding remarks.

In this Chapter, I took into account the theory of Predictive Coding in three versions: Hohwy's one, Clark's Predictive Processing, and the reassessments of those proposals according to an enactive perspective.

My discussion was aimed at understanding if Clark's Action-Oriented Predictive Processing is able to declare peace between the enactive approach to cognition and his EMH (Clark 2015a). My exam of Clark's proposal showed that peace cannot be declared yet. Indeed, Clark's Action Oriented Predictive Processing still makes use of the notion of internal representation (AOR) to explain action-perception loops. In an enactive perspective, the use of this epistemic posit is unproductive: AORs do not explain action-perception loops, but, on the contrary, they conceal the very role of embodied, situated action. To develop a theory of the "situated, active and predictive mind", the explanation has to focus on the very role of action, conceived as an external process that involves an embodied, skillful agent and her Umwelt, defined as a rich field affordances that triggers patterns of action.

Crucially, this dynamics is thought to be affectively connoted. This point is particularly interesting because it not only gives a more multilayered account of cognitive process than Clark's one does, in which the topic of affectivity is not discussed in detail, but it also offers the conceptual tools for an explanatory framework that does not rely on representations. At the same time, the enactive approach to PP does not state that cognitive process are just mere reactions to environmental conditions.

This is a crucial point for the debate developed in cognitive sciences. Indeed, sometimes it has been objected that the enactive approach to cognition is a new form of behaviorism (Block 2001) because of the emphasis that this theory puts on the sensorimotor dimension of cognitive processes, which are said not being mediated by computations and representations. The lack of this representational mediation is said to entail that perceptual and active cognitive processes are conceived in terms of an input-output dynamics.

The way the enactive approach to cognition accounts for action-perception loops is able to avoid this scenario (Bruineberg, Rietveld 2014), sometimes called "desert landscape scenario" (Clark 2013a) to emphasize how the enactive theory avoids internal representations and goals to explain action-perception loops and anticipatory processes. Indeed, by situating affectivity in "the in-between" the agent and the environment, the enactive approach to cognition accounts for the relation between the two components of the coupled system in a mediated but non-representational way. Affect mediates between

the organism and the environment, and it does so in an active way, namely by triggering patterns of action that shape the agent's states of action-readiness, crucial for anticipatory aspects of cognitive processes.

Conclusion and future work

Where does the mind stop and the rest of the world begin? This is the challenging question Clark and Chalmers' revolutionary article the "Extended Mind" (Clark, Chalmers 1998) begins with.

My work mostly dealt with this question. I started to consider the problem of the mind's borders in my exam of some contemporary approaches to cognition (Cognitive Linguistics, Grounded Cognition, Embedded, Enacted and Extended Cognition). I pointed out how those theories try to extend the mind's borders by rethinking of the cognitive role of the body, action, and the environment in cognitive practices.

I claimed that, in the 4Es' debate, the Extended Mind Hypothesis and the Enactive approach to cognition are the most interesting theories we should look at to conceive of the mind as "extended" or "extensive". Indeed both endorse an ontological claim about the mind. According to the EMH, the mind is constituted by the manipulation of information embedded in the brain or in externalities (e.g. artifacts) distributed in the environment. Also according to the enactive approach to cognition the mind does not stop at the boundaries defined by the skull: the mind is embodied and enacted; it is a process that develops through the active and affective relation between the agent and her environment.

Despite both theories endorse revolutionary claims about the nature of the mind, they do so by following different philosophical strategies. I claimed that this influences the success of these two philosophical projects.

The EMH endorses an extended form of functionalism. Even if at first glance extended functionalism sounds like an appealing concept for the project of "mind-extension", it does not actually succeed in radically reassessing the concept of mind. Indeed, this philosophical framework entails a locational claim about cognition, and it implicitly gives a cognitive primacy to internal cognitive resources, which are considered to be representational states. The EMH's "extended mind" seems to still epistemically stop at the boundaries of the skull: representational states are the model Clark looks at in order to think of the cognitive role of the body and the world. Moreover, this explanation is not as active as it should be: the theory is grounded on the concept of mental state and, by explaining action through internal representations, it underplays the cognitive role of world-engaging embodied action.

On the contrary, the enactive approach to cognition gives a relational and more active account of cognitive processes, and then of the mind. Embodied, world-engaging action is central in this explanation, and crucially it is not considered to be dependent on or forerun by internal representational states. This makes the "extensive" account of the mind offered by the enactive approach to cognition stronger than Clark's EMH. Indeed, the enactive approach to cognition gives a dynamical and processual explanation of cognitive processes. The emphasis the enactive approach to cognition puts on embodied action -conceived as the *attunement* between the agent and the environment- allows the explanation to overcome the problems the EMH's locational and representational commitment entails. Those problems are: i) a static account of the mind, which does not explain *how* the mind extends to externalities, ii) an internalistic prejudice, according to which there is a gap between the mind and the world that should be filled by representations.

Even if the EMH tries to redefine the concept of representation in order to make it active and dynamical -and then suitable to account for the mind as an extended process- I claimed that this explanation is not successful at all. In fact, Clark's AORs are philosophically unwarranted, because they do not fulfill representational criteria. Moreover, the way Clark defines those representations does not account for action. Like representations postulated by first-wave cognitive sciences, AORs are internal items that anticipate and control action, they are not endowed with *intrinsic active* features, others concepts of representation -such as Peirce's signs- are endowed with.

Therefore, I claimed that an "extensive" approach to cognition should not buy AORs: AORs do not play any interesting job in explaining cognitive processes (e.g. embodied and enacted selective responses to the environment, anticipation of action-perception loops) as active.

That is why I endorsed an enactive perspective on cognition. This one avoids AORs, then it does not entail some of the problems the EMH suffers from. Moreover, even if the enactive approach to cognition rejects internal representations, it does not explain cognitive processes as mere reactions to inputs. Indeed, the enactive approach to cognition emphasizes the role of affects in cognitive processes a lot. Affect -which is considered to be embodied and situated- does the same job internal representations are supposed to do: it contributes to *adaptive* dynamics, selective responsiveness to stimuli and anticipation of action perception-loops are expressive of. Crucially, it does so avoiding rhetorical slips AORs-based explanations entail; furthermore, it does not give an epistemic primacy to internal cognitive resources.

This not only makes the enactive approach to cognition more suitable to think of the mind as a complex, extensive and enactive process, but it also makes the theory not liable to objections according to which non-representational accounts of cognition are new forms of behaviorism (Aizawa 2014; Block 2001; Carruthers 2015; Shapiro 2011, also cited in Gallagher 2017: 132). Indeed, by means of affectively driven action, the enactive approach to cognition explains how cognitive practices build up the coupled cognitive system made of the cognitive agent and her Umwelt. In this system, the agent's cognitive dynamics and the environmental ones are attuned the one with the other. There are not two different systems -the agent and the environment- and there is not one system (the agent or the organism) that takes the other system as its input. That is why the enactive approach to cognition cannot be considered to be a new form of behaviorism. Cognitive practices unfold through the relation between parts of the same system. This relation, as explained in Chapter III, is not thought to be a mere reaction to inputs, but it is said to develop in an affective way. Indeed, affects not only structure the coupled cognitive system, but also account for the interaction between parts of the system. This interaction, I claim, is more direct than the one described by representational approaches to cognition, which postulate a complex cognitive slice between the agent and the environment. Nevertheless, it entails a kind of affective mediation, affective and embodied tension is expressive of: this distinguishes mechanical behavior from cognitive experience.

Given this theoretical importance of affect in the enactive approach to cognition -crucial to understand the relation between enactivism and other theories, such as minimal representational approaches to cognition like the EMH, and fundamental to defend enactivism from its critiques- it could be interesting to devote my future work to more detailed considerations about the role of affectivity in cognitive experience.

This would also help me to make my current work more complete. Indeed, here I focused on specific kinds of cognitive processes, perception and action, and I always took into account an agent and her environment. A more accurate work on the affective dimension of cognition would help me to consider social cognition and collective action as well: the individual agent and the perceptual environment of this work would become a plural agent and an intersubjective environment.

Works cited

Adams, F., Aizawa, K. (2001), "The bounds of cognition", *Philosophical Psychology,* 14, pp. 43-64.

Adams, F., Aizawa, K. (2009), "Why the mind is still in the head", in P. Robbins, M. Aydede [Eds.], *Cambridge Handbook of Situated Cognition*, Cambridge, Cambridge University Press, pp. 78-95.

Adams, F., Aizawa, K. (2010a), *The bounds of cognition*, Chichester, Wiley-Blackwell.

Adams, F., Aizawa, K. (2010b), "Defending the bounds of cognition", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 67-80.

Agre, P. E. (1997), *Computation and human experience*, Oxford, Oxford University Press.

Aizawa, K. (2010a), "The coupling-constitution fallacy revisited", *Cognitive Systems Research*, 11, pp. 332-342.

Aizawa, K. (2010b), "Computation in cognitive science: it is not all about Turing-equivalent computation", *Studies in History and Philosophy of Science*, 41, pp. 227-236.

Aizawa, K. (2014), "The Enactivist Revolution", *Avant*, 2, pp. 19-42.

Anderson, M. L. (2003), "Embodied Cognition: a Field Guide", *Artificial Intelligence*, 149, pp. 91-130.

Anderson, M. L. (2010), "Neural reuse: A fundamental organizational principle of the brain", *Behavioral and Brain Sciences*, 33, pp. 245 -311.

Atã, P., Queiroz, J. (2014), "Iconicity in Peirce's situated Semiotics", in T. Thellefsen, B. Sørensenm [Eds.], *Peirce in his own words. 100 years of semiotics, communication, and cognition*, Berlin, Mouton de Gruyter, pp. 283-290.

Atã, P., Queiroz, J. (2016), "Habit in Semiosis: Two Different Perspectives Based on Hierarchical Multi-level System Modeling and Niche Construction Theory", in D. West, M. Anderson [Eds.], *Consensus on Peirce's concept of habit. Before and beyond consciousness*, Berlin, Springer, pp.109-122.

Aydin, C. (2015), "The artifactual mind: overcoming the 'inside–outside' dualism in the extended mind thesis and recognizing the technological dimension of cognition"*, Phenomenology and the Cognitive Sciences*, 14 (1), pp. 73-94.

Barandiaran, X., Di Paolo, E. (2014), "A genealogical map of the concept of habit", *Frontiers in Human Neuroscience*, 8, pp. 1-7.

Barandiaran, X. E., Egbert, M. D. (2014), "Norm-Establishing and Norm-Following in Autonomous Agency", *Artificial Life*, 20 (1), pp. 5-28.

Barandiaran, X. E. (2016), "Autonomy and Enactivism. Towards a theory of sensorimotor autonomous agency", *Topoi,* pp. 1-22.

Barrett, L., Bar, M. (2009), "See it with feeling: affective predictions during object perception", *Philosophical Transactions of the Royal Society*, 364, pp. 1325-1334.

Barsalou, L. (1999), "Perceptual Symbols systems", *Behavioral and Brain Sciences*, 22, pp. 577-609.

Barsalou, L., Prinz, J. (2000), "Steering a course for embodied representation", in E. Dietrich, A. Markman [Eds.], *Cognitive Dynamics: Conceptual Change in Humans and Machines*, Mahwah, Lawrence Erlbaum, pp. 51-77.

Barsalou, L., Simmons, W. K, Barbey, A. K., Wilson, C. D. (2003), "Grounding conceptual knowledge in modality-specific systems", *Trends in Cognitive Sciences*, 7 (3), pp. 84-91.

Barsalou, L. (2008), "Grounded Cognition", *Annual Review of Psychology*, 59, pp. 617-645.

Barsalou, L. (2009), "Simulation, situated conceptualization, and prediction", *Philosophical Transactions of the Royal Society*, 364, pp. 1281-1289.

Barsalou, L. (2016), "On Staying Grounded and Avoiding Quixotic Dead Ends", *Pshyconomic Bulletin and Review*, 23 (4), pp. 1122-1142.

Bechara, A., Damasio, H., Tranel, D., Damasio, A. R. (1997), "Deciding advantageously before knowing the advantageous strategy", *Science*, 275 (5304), pp. 1293 – 1295.

Block, N. (2001), "Behaviorism Revisited", *Behavioral and Brain Sciences*, 24 (5), pp. 977–978.

Blokpoel, M., Kwisthout, J., Rooij, I. (2012), "When can predictive brains be truly Bayesian?", *Frontiers in Psychology*, 3, pp. 1-3.

Brooks, R. (1989), "Fast, cheap and out of control. A robot invasion of the solar system", *Journal of The British Interplanetary Society,* 42*,* pp. 478-485.

Brooks, R. (1990), "Elephants don't play chess", *Robotics and Autonomous Systems*, 6 (1-2), pp. 3-15.

Brooks, R. (1991), "Intelligence without representations", *Artificial Intelligence*, 47, pp. 139-159.

Brooks, R. (1995), "Intelligence without reason", *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, Sydney, Australia (August 1991), pp. 569-595.

Bruineberg, J., Rietveld, E. (2014), "Self-organization, free energy minimization, and optimal grip on a field of affordances", *Frontiers in Human Neuroscience*, 8, pp. 1-14.

Bruineberg, J., Kiverstein, J., Rietveld, E. (2016), "The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective", *Synthese*, Online First, pp. 1-28.

Burge, T. (1979), "Individualism and the Mental," in P. French, T. Uehling, H. Wettstein [Eds.], *Midwest Studies in Philosophy*, IV, Minneapolis, University of Minnesota Press, pp. 73–121.

Burke, T. (2008), "(Anti)Realist Implications of a Pragmatist Dual-Process Active-Externalist Theory of Experience", *Philosophia Scientiae*, 12 (1), pp. 187-211.

Burke, T. (2014), "Extended Mind and Representation", in J. R. Shook, T. Solymosi [Eds.], *Pragmatist Neurophilosophy. American Philosophy and the Brain*, New York/London, Bloomsbury Publishing, pp. 177-202.

Calcaterra, R. M. (2006), "Psychology and Anti-psychologism in Peirce", in R. Fabbrichesi, S. Marietti [Eds.], *Semiotics and Philosophy in Charles Sanders Peirce,* Cambridge, Cambridge Scholars Press, pp. 35-50.

Cappuccio, M., Wheeler, M. (2012), "Ground-Level Intelligence: Action-Oriented Representation and the Dynamics of the Background", in Z. Radman [Ed.], *Knowing without Thinking. Mind, Action, Cognition and the Phenomenon of the Background*, New York, Palgrave McMillan, pp. 13-36.

Cappuccio, M., Froese, T. (2014), "Introduction to *Making Sense of Non-Sense*", in M. Cappuccio, T. Froese [Eds.], *Enactive Cognition at the Edge of Sense-Making. Making Sense of Non-Sense,* New York, Palgrave McMillan, pp. 1-38.

Carruthers, P. (2015), "Perceiving mental states", *Consciousness and Cognition*, 36, pp. 498-507.

Casasanto, D. (2009), "Embodiment of abstract concepts: good and bad in right and left-handers", *Journal of Experimental Psychology: General*, 138, pp. 351-367.

Casasanto, D. (2011), "Different Bodies, Different Minds: The Body Specificity of Language and Thought", *Current Directions in Psychological Science*, 20 (6), pp. 378 –383.

Casasanto, D., Chrysikou, E. G. (2011), "When left is 'right': Motor fluency shake abstract concepts", *Psychological Science*, 22, pp. 419-422.

Casasanto, D. (2014), "Bodily relativity", in L., A. Shapiro [Ed.], *The Routledge Handbook of Embodied Cognition*, London/New York, Routledge, pp. 108 – 117.

Chemero, A. (1998), "A Stroll Through the Worlds of Animats and Humans: Review of *Being There: Putting Brain, Body and World Together Again* by Andy Clark", *Psyche*, 4, On Line Only.

Chemero, A., Eck, D. (1999), "An exploration of representational complexity via coupled oscillator systems", in U. Priss [Ed.], *Proceedings of the 1999 Midwest AI and Cognitive Science Conference*, Cambridge, AAAI Press.
Chemero, A, (2000), "Representation and reliable presence", *Conceptus Studien,* 14, pp. 9-25.
Chemero, A. (2003), "An outline of a theory of affordances", *Ecological Psychology*, 15 (2), pp. 181–195.
Chemero, A. (2009), *Radical Embodied Cognitive Science*, Cambridge (MA), The MIT Press.
Chiel, H., Beer, R. (1997), "The brain has a body. Adaptive behavior emerges from interactions of nervous system, body and environment", *Trends in Neuroscience*, 20, pp. 553-57.
Chomsky, N. (1959), "A Review of B. F. Skinner's *Verbal Behavior"*, *Language,* 35 (1), pp. 26-58.
Chomsky, N. (1965), *Aspects of the Theory of Syntax*, Cambridge (MA), The MIT Press.
Chow, S. J. (2013), "What's the problem with the frame problem?", *Review of Philosophy and Psychology*, 4 (2), pp. 309-331.
Cisek, P. (2007), "Cortical mechanisms of action selection: The affordance competition hypothesis", *Philosophical Transactions of the Royal Society*, B (362), pp. 1585-1599.
Cisek, P., Kalaska, J. F. (2011), "Neural mechanisms for interacting with a world full of action choices", *Annual Review of Neuroscience*, 33, pp. 269-298.
Clark, A., Toribio, J. (1994), "Doing without representing?", *Synthese*, 101, pp. 401-431.
Clark, A. (1995), "Moving minds: Re-thinking representation in the heat of situated action", *Philosophical Perspectives*, 9, pp. 89-104.
Clark, A. (1997a), *Being there. Putting Brain, Body, and World Together Again*, Cambridge (MA), The MIT Press.
Clark, A. (1997b), "The dynamical challenge", *Cognitive Science*, 21 (4), pp. 461-481.
Clark, A., Chalmers, D. (1998), "The Extended Mind", *Analysis,* 58 (1), pp.7-19.
Clark, A. (1998), "Magic Words: how language augments human computation", in P. Carruthers, J. Boucher [Eds.], *Language and thought: interdisciplinary themes*, Cambridge, Cambridge University Press, pp. 162-183.
Clark, A. (1999), "An Embodied Cognitive Science?", *Trends in Cognitive Sciences*, 3 (9), pp. 345-351.

Clark, A., Grush, R. (1999), "Towards a Cognitive Robotics", *Adaptive Behavior*, 7 (1), pp. 5-16.
Clark, A. (2001), "Natural-Born Cyborgs?", in M. Beynon, C.L. Nehaniv, K. Dautenhahn [Eds.], *Cognitive Technology: Instruments of Mind*, Berlin, Springer, pp. 17-24.
Clark, A. (2003), *Natural-Born Cyborgs: Technologies and the Future of Human Intelligence*, Oxford, Oxford University Press.
Clark, A. (2005a), "Word, Niche, Super-Niche: How Language Makes Mind Matter More", *Theoria*, 20, pp. 255-268.
Clark, A. (2005b), "Intrinsic content, active memory and the Extended Mind", *Analysis*, 65, pp. 1-11.
Clark, A. (2006), "Cognitive Complexity and the Sensorimotor Frontier", *Aristotelian Society Supplementary Volume*, 80 (1), pp. 43-65.
Clark, A. (2008), *Supersizing the Mind. Embodiment, Action and Cognitive Extension*, Oxford, Oxford University Press.
Clark, A., Kiverstein, J. (2009), "Mind Embodied, Embedded, Enacted: One Church or Many*?", Topoi*, 28, pp. 1–7.
Clark, A. (2010a), "Coupling, Constitution, and the Cognitive Kind: A Reply to Adams and Aizawa", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 81-100.
Clark, A. (2010b), "Memento's revenge: The Extended Mind, Extended", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 43-66.
Clark, A. (2012), "Dreaming the Whole Cat: Generative Models, Predictive Processing, and the Enactivist Conception of Perceptual Experience", *Mind*, 121 (483), pp. 753-771.
Clark, A. (2013a), "Whatever the next? Predictive brains, situated agents and the future of cognitive science", *Behavioral and Brain Sciences*, 36 (3), pp. 181-204.
Clark, A. (2013b), "The many faces of precision. Replies to commentaries on 'Whatever the next? Predictive brains, situated agents and the future of cognitive science'", *Frontiers in Psychology*, 4, pp. 1-9.
Clark, A. (2015a), "Predicting Peace: The End of the Representation Wars. A Reply to Michael Madary", in T. Metzinger, J. M. Windt [Eds.], Open MIND, 7 (R), pp.1-7, doi: 10.15502/9783958570979.
Clark, A. (2015b), "Embodied Prediction", in T. Metzinger, J. M. Windt [Eds.], Open MIND, 7 (R), pp.1-21, doi: 10.15502/9783958570115.
Clark, A. (2015c), "Radical Predictive Processing", *The Southern Journal of Philosophy*, 53, Spindel Supplement, pp. 3-27.

Clark, A. (2016), *Surfing Uncertainty. Prediction, Action, and the Embodied Mind,* Oxford/New York, Oxford University Press.

Clark, A. (2017), "How to Knit Your Own Markov Blanket: Resisting the Second Law with Metamorphic Minds", in T. Metzinger & W. Wiese [Eds.], *Philosophy and Predictive Processing*: 3, Frankfurt am Main: MIND Group, pp. 1-19, doi: 10.15502/9783958573031.

Clowes, R.W, Mendonça, D. (2016), "Representation Redux: Is there still a useful role for representation to play in the context of embodied, dynamicist and situated theories of mind?", *New Ideas in Psychology*, 40, pp. 26-47.

Colapietro, V. (1989), *Peirce's Approach to the Self: A Semiotic Approach on Human Subjectivity,* Albany NY, State University of New York Press.

Colombetti, G. (2007), "Enactive appraisal", *Phenomenology and the Cognitive Sciences*, 6, pp. 527–546.

Colombetti, G. (2014), *The Feeling Body. Affective Science Meets the Enactive Mind*, Cambridge (MA), The MIT Press.

Colombetti, G. (2015), "Enactive Affectivity, Extended", *Topoi*, 34 (1), pp. 1-11.

Daddesio, T. (1995*), On Minds and Symbols. The Relevance of Cognitive Science for Semiotics*, Berlin/New York, Mouton de Gruyter.

Damasio, A. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain,* New York, G. P. Putnam.

Dawson, M. (2014), "Embedded and Situated Cognition", in L. A. Shapiro [Ed.], *The Routledge Handbook of Embodied Cognition*, London/New York, Routledge.

De Jaegher, H., Di Paolo (2007), "Participatory sense-making. An enactive approach to social cognition", *Phenomenology and the Cognitive Sciences*, 6, pp. 485- 507.

De Jaegher, H., Di Paolo, Gallagher, S. (2010), "Can social interaction constitute social cognition?", *Trends in Cognitive Sciences*, 14 (10), pp. 441-447.

De la Vega, I., Dudschig, C., De Filippis, M., Lachmair, M., Kaup, B. (2013), "Keep your hands crossed: The valence-by-left/right interaction is related to hand, not side, in an incongruent hand –response key assignment", *Acta Psychologica,* 142 (2), pp. 273-277.

Deacon, T. (2014), "Semiosis: from taxonomy to process", in T. Thellefsen, B. Sørensen [Eds.], *Peirce in his own words. 100 years of semiotics, communication, and cognition*, Berlin, Mouton de Gruyter, pp. 95-104.

De Jesus, P. (2016), "From enactive phenomenology to biosemiotic enactivism", *Adaptive Behavior*, 24 (2), pp. 130-146.

Degenaar, J., Myin, E. (2014), "Representation-hunger reconsidered", *Synthese*, 191, pp. 3639-3648.

Degenaar, J., O' Regan, K. (2015), "Sensorimotor Theory and Enactivism", *Topoi* , pp. 1-15.

Descombes, V. (2001), *The Mind's Provisions. A critique to Cognitivism*, Princeton, Oxford University Press.

Dewey, J. (1938), *Logic. The Theory of Inquiry*, New York, Henry Holt and Company.

Di Paolo, E. (2005), "Autopoiesis, adaptivity, teleology, agency", *Phenomenology and the Cognitive Sciences*, 4, pp. 429-452.

Di Paolo, E. (2009), "Extended Life", *Topoi*, 28 (1), pp. 9-21.

Di Paolo, E., De Jaegher, H. (2012), "The interactive brain hypothesis", *Frontiers in Human Neuroscience*, 6, pp. 1-16.

Donald, M. (2010), "The Exographic Revolution: Neuropsychological Sequelae", in L. Malafouris, C. Renfrew [Eds.], *The cognitive life of things. Recasting the boundaries of the mind*, Cambridge, McDonald Institute for Archaeological Research, pp. 71-80.

Dreyfus, H. L. (1972), *What Computers Can't Do: The Limits of Artificial Reason*, New York, Harpers & Row Publishers.

Dreyfus, H. L. (1995), *Being-in-the World. A commentary on Heidegger's Being and Time, Division I*, Cambridge (MA), The MIT Press.

Dreyfus, H. L. (2002a), "Refocusing the question: can there be skillful coping without propositional representations or brain representations?", *Phenomenology and the Cognitive Sciences*, 1 (4), pp. 413-425.

Dreyfus, H. L. (2002b), "Intelligence without representation: Merleau-Ponty's critique of mental representation", *Phenomenology and the Cognitive Sciences,* 1 (4), pp. 367-383.

Dreyfus, H. L. (2005), "Merleau-Ponty and Recent Cognitive Science", in T. Carman, M. N. Hansen [Eds.], *The Cambridge Companion to Merleau-Ponty*, Cambridge, Cambridge University Press, pp. 129-150.

Dreyfus, H. L, Kelly, S. D. (2007), "Heterophenomenology: Heavy-handed sleight-of-hand", *Phenomenology and the Cognitive Sciences*, 6 (1-2), pp. 45-55.

Dreyfus, H. L. (2008), "Why Heideggerian AI failed and how fixing it would require making it more Heideggerian", in P. Husbands, M. Wheeler [Eds.], *The Mechanical Mind in History*, Cambridge (MA), The MIT Press, pp. 331-371.

Eco, U. (1968), *La struttura assente. La ricerca semiotica e il metodo strutturale*, Milano, Bompiani.

Eco, U. (1975), *Trattato di Semiotica generale*, Milano, Bompiani/ *A Theory of Semiotics*, Bloomington, Indiana University Press, 1976.
Eco, U. (1976), "Peirce's notion of Interpretant", *MLN*, 91 (6), pp. 1457-1472.
Eco, U. (1984), "Metaphor, Dictionary, Encyclopedia", *New Literary History*, 15 (2), pp. 255-271.
Engel, A.K., Maye, A., Kurthen, M., Köning, P. (2013), "Where's the action? The pragmatic turn in cognitive science", *Trends in Cognitive Science*, 17 (5), pp. 202-209.
Fabbrichesi, R. (1986), *Sulle tracce del segno. Semiotica, faneroscopia e cosmologia nel pensiero di Charles S. Peirce*, Firenze, La Nuova Italia.
Fabbrichesi, R., Leoni, F. (2005), *Continuità e variazione. Leibniz, Goethe, Peirce, Wittgenstein con un'incursione kantiana*, Milano, Mimesis.
Fabbrichesi, R. (2017), "The Iconic Ground of Gestures. Peirce, Wittgenstein, and Foucault", in K. A. Hull, R. K. Atkins [Eds.], *Peirce on Perception and Reasoning. From Icons to Logic*, New York/London, Routledge, pp. 54-60.
Farina, A., Belgrano, A. (2005), "The eco-field hypothesis: toward a cognitive landscape", *Landscape Ecology*, pp. 5-17.
Farina, A. (2008), "The Landscape as a Semiotic Interface between Organism and Resources", *Biosemiotics*, 1, pp. 75-83.
Fisch, M. H. (1986), *Peirce, Semeiotic, and Pragmatism. Essays by Max H. Fisch,* K. L. Ketner, C. J. W. Kloesel [Eds.], Bloomington/Indianapolis, Indiana University Press.
Fodor, J. (1975), *The Language of Thought,* New York, Thomas Y. Crowell Company.
Fodor, J. (1983) *The Modularity of Mind. An Essay on Faculty Psychology,* Cambridge (MA), The MIT Press.
Fodor, J. (1987), *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge (MA), The MIT Press.
Fodor, J. (2008), *LOT 2. The Language of Thought Revisited*, Oxford, Oxford University Press.
Foucault, M. (1969), *L'Archéologie du Savoir*, Paris, Gallimard.
Frijda, N. H., Ridderinkhof, R. K., Rietveld, E. (2014), "Impulsive action: emotional impulses and their control", *Frontiers in Psychology*, 5 (518), Online.
Friston, K., Stephan, K. E. (2007), "Free-energy and the brain", *Synthese*, 159, pp. 417–458.
Friston, K. (2011a), "Prediction, perception and agency", *International Journal of Psychophysiology*, 83, pp. 248-252.
Friston, K. (2011b), "Embodied Inference: or 'I think therefore I am, if I am what I think", in W. Tschacher, C. Bergomi [Eds.], *The implications of*

embodiment (Cognition and Communication), Exeter, Imprint Academic, pp. 89-125.

Friston, K. (2012), "The history of the future of the Bayesian brain", *NeuroImage*, 62 (2), pp. 1230-1233.

Friston, K. (2013), "Life as we know it", *Journal of the Royal Society, Interface*, 10, DOI: 10.1098/rsif.2013.0475.

Fumagalli, A. (2005), *Il reale nel linguaggio. Indicalità e realismo nella semiotica di Peirce*, Milano, Vita e Pensiero, Pubblicazioni dell'Università Cattolica del Sacro Cuore.

Fusaroli, R., Granelli, T., Paolucci, C. (2011) [Eds.], *The External Mind Perspectives on Semiosis, Distribution and Situation in Cognition,* Versus 112-113, Quaderni di studi semiotici,

Gallagher, S. (1986), "Body Image and Body Schema: A Conceptual Clarification", *Journal of Mind and Behavior*, 7 (4), pp. 541-554.

Gallagher, S. (2005), *How the Body Shapes the Mind*, Oxford, Oxford University Press.

Gallagher, S. (2008a), "Are minimal representations still representations*?", International Journal of Philosophical Studies*, 16 (3), pp. 351-369.

Gallagher, S. (2008b), "Cognitive Science", in R. Diprose, J. Reynolds [Eds.], *Merleau-Ponty: Key Concepts,* Stocksfield, Acumen, pp. 207-217.

Gallagher, S., Zahavi, D. (2008), *The Phenomenological Mind. An Introduction to Philosophy of Mind and Cognitive Science*, New York, Routledge.

Gallagher, S. (2009), "Philosophical antecedents of situated cognition", in P. Robbins, M. Aydede [Eds.], *Cambridge Handbook of Situated Cognition*, Cambridge, Cambridge University Press, pp. 35-51.

Gallagher, S. (2011a), "Interpretations of embodied cognition", in W. Tschacher, C. Bergomi [Eds.], *The Implications of Embodiment: Cognition and Communication*, Exeter, Imprint Academic, pp. 59-70.

Gallagher, S. (2011b), "The Overextended Mind", *Versus-Quaderno di Studi Semiotici*, 112-113, pp. 57-68.

Gallagher, S. (2012), "On the Possibility of Naturalizing Phenomenology", in D. Zahavi [Ed.], *The Oxford Handbook of Contemporary Phenomenology*, Oxford, Oxford University Press, pp. 70-93.

Gallagher, S., Miyahara, K. (2012), "Neo-Pragmatism and Enactive Intentionality", in J. Schulkin [Ed.], *Action, Perception and the Brain. Adaptation and Cephalic Expression,* Basingstoke, Hampshire, Palgrave McMillan, pp. 117-146.

Gallagher, S. (2013), "The socially extended mind", *Cognitive system Research*, 25-26, pp. 4-12.

Gallagher, S. (2014), "Pragmatic interventions into enactive and extended conceptions of cognition", *Philosophical Issues*, 24, pp. 110-126.

Gallagher, S., Bower, M. (2014), "Making enactivism even more embodied", *Avant*, 5 (2), pp. 232-247.

Gallagher, S., Ransom, T., G. (2015), "Artifacting Minds: Material Engagement Theory and Joint Action", in G. Etzelmüller, C. Tewes [Eds.], *Embodiment in Evolution and Culture*, Tübingen, Mohr Siebeck, pp. 337-352.

Gallagher, S., Allen, M. (2016), "Active inference, enactivism and the hermeneutics of social cognition", *Synthese*, Online Frist, pp. 1-22.

Gallagher, S. (2017), *Enactivist Interventions. Rethinking the Mind*, Oxford, Oxford University Press.

Gardner, H. (1985), *The Mind's New Science. A history of the cognitive revolution*, New York, Basic Books.

Geeraerts, D., Cuyckens, H. [Eds.] (2007), *The Oxford Handbook of Cognitive Linguistics*, Oxford, Oxford University Press.

van Gelder, T. (1995), "What might cognition be, if not computation?", *Journal of Philosophy*, 91, pp. 345–81.

Gibbs, R. W., Colston, H., L. (1995), "The cognitive psychological reality of image schemas and their transformations", *Cognitive Linguistics*, 6 (4), pp. 347-378.

Gibson, J. (1979), *The ecological approach to visual perception*, Boston, Houghton Mifflin/ Classic edition published by New York/London, Psychology Press, 2015.

Gładziejewski, P. (2016), "Predictive Coding and Representationalism", *Synthese*, 193 (2), pp. 559-582.

Glenberg, A., M. (1997), "What memory is for", *Behavioral and Brain Sciences*, 20, pp. 1-55.

Glenberg, A., M., Robertson D., A. (2000), "Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning", *Journal of Memory and Language*, 43, pp. 379-401.

Glenberg, A., M., Kaschak, M. P. (2002), "Grounding language in action", *Psychonomic Bulletin and Review*, 9 (3), pp. 558-565.

Goldman, A. (2006), *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*, Oxford, Oxford University Press.

Grush, R. (2004), "The emulation theory of representation: Motor control, imagery, and perception", *Behavioral and Brain Sciences*, 27, pp. 377–442.

Hardwick, C., S. [Ed.] (1977), *Semiotics and Significs: The correspondace between Charles S. Peirce and Victoria Lady Welby,* Bloomington, Indiana University Press.

Harnad, S. (1990), "The Symbol Grounding Problem", *Physica*, 42, pp. 335-346.

Haselager, W. F. G, Van Rappard, J. F. H. (1998), "Connectionism, Systematicity, and the Frame Problem", *Minds and Machines,* 8, pp. 161–179.

Haugeland, J. (1991), "Representational genera", in W. Ramsey, S. Stich, J. Garron, [Eds.], *Philosophy and Connectionist Theory*, Hillsdale, Lawrence Erlbaum, pp. 171-206.

Heidegger, M. (1922), *Phänomenologische Interpretationen zu Aristoteles*, Reclam, Stuttgart, 2003/ *Phenomenological Interpretations of Aristotle*, English translation by R. Rojcewicz, Bloomington, Indiana University Press, 2001.

Heidegger, M. (1927), *Sein und Zeit*, Halle, Niemeyer Verlag / *Being and Time*, English translation by J. Macquarrie and E. Robinson, Oxford, Blackwell, 1962.

Heilbron, M., Chait, M. (2017), "Great expectations: is there evidence for Predictive Coding in auditory cortex?", *Neuroscience*, in press, https://doi.org/10.1016/j.neuroscience.2017.07.061.

von Helmholtz, H. (1866), "Concerning the perceptions in general", in *Treatise on Physiological Optics*, Vol. III, English translation by J. P. C. Southall, New York, Dover, 1962.

Hjelmslev, L. T. (1943) *Omkring sprogteoriens grunglaeggelse,* Copenhagen, Munksgaard. / *Prolegomena to a Theory of Language* (1969)*,* English translation by F. J. Whitfield, Madison/Milwaukee/London, The University of Wisconsin Press.

Hoffmeyer, J. (2015), "Introduction: Semiotic Scaffolding", *Biosemiotics*, 8 (2), pp. 153-158.

Holroyd, C., Coles, M. (2002), "The neural basis of human error processing. Reinforcement learning, dopamine, and the error-related negativity", *Psychological Review*, 109, pp. 679-709.

Hohwy, J. (2012), "Attention and conscious perception in the hypothesis testing brain", *Frontiers in Psychology*, 3, pp. 1-14.

Hohwy, J. (2013), *The Predictive Mind*, Oxford, Oxford University Press.

Hohwy, J. (2014), "The Self-Evidencing Brain", *Noûs*, 50 (2), pp. 1-27.

Hohwy, J. (2015), "The Neural Organ Explains the Mind", in T. Metzinger & J. M. Windt [Eds.], Open MIND, 19 (T), Frankfurt am Main: MIND Group, pp. 1-37, doi: 10.15502/9783958570016.

Hohwy, J. (2017), "How to Entrain Your Evil Demon", in T. Metzinger & W. Wiese [Eds.], *Philosophy and Predictive Processing*: 2, Frankfurt am Main: MIND Group, pp. 1-15, doi: 10.15502/9783958573048.

Hufendiek, R. (2015), *Embodied Emotions: A Natural Approach to a Normative Phenomenon,* New York/London, Routledge.

Hufendiek, R. (2017), "Affordances and the Normativity of Emotions", *Synthese*, 194, pp. 4455–4476.

Hunter, J. F. M. (1968), "'Forms of Life' in Wittgenstein's 'Philosophical Investigations'", *American Philosophical Quarterly*, 5 (4), pp. 233-243.

Hurley, S. (1998), *Consciousness in Action*, Cambridge (MA), Harvard University Press.

Hurley, S. (2010), "The varieties of externalism", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 101-154.

Hutto, D. (2006), "Unprincipled engagements. Emotional experience, expression and response", in R. Menary [Ed.], *Radical Enactivism. Intentionality, Phenomenology and Narrative. Focus on the philosophy of Daniel Hutto*, Amsterdam/Philadelphia, John Benjamins Publishing Company, pp. 13-38.

Hutto, D. (2011), "Philosophy of Mind's New Lease on Life: Autopoietic Enactivism meets Teleosemiotics", *Journal of Consciousness Studies,* 18 (5-6), pp. 44-64.

Hutto, D. (2012), "Truly Enactive Emotion", *Emotion Review*, 4 (2), pp. 176-181.

Hutto, D. (2013), "Exorcising action oriented representations: ridding cognitive science of its Nazgûl", *Adaptive Behavior*, 21 (3), pp. 142-150.

Hutto, D. D, Myin, E. (2013), *Radicalizing Enactivism. Basic Minds without content*, Cambridge (MA), The MIT Press.

Hutto, D. D., Kirchhoff, M. D., Myin, E. (2014), "Extensive enactivism: Why keep it all in?", *Frontiers in human neuroscience*, 8, pp. 1-11.

Hutto, D. (2015), "Overly Enactive Imagination? Radically Re-Imagining Imagining", *The Southern Journal of Philosophy*, 53, Spindel Supplement, pp. 68-89.

Hutto, D. (2017), "Getting into predictive processing's great guessing game: Bootstrap heaven or hell?", *Synthese*, Online First, pp. 1-14.

Hutto, D., Myin, E. (2017), *Evolving Enactivism. Basic minds meet content,* Cambridge (MA), The MIT Press.

Ingold, T. (2000), *The Perception of the Environment. Essays on livelihood, dwelling and skill*, London/New York, Routledge.

Johnson, M. (2006), "Cognitive Science", in J. R. Shook, J. Margolis [Eds.], *A Companion to Pragmatism*, Oxford, Blackwell, pp. 369-377.

Johnson, M., Rohrer, T. (2007), "We are living creatures: Embodiment, American Pragmatism, and the Cognitive Organism", in J. Zlatev, T. Ziemke,

R. Frank, R. Dirven [Eds.], *Body, Language, and Mind*, Vol. 1, Berlin, Mouton de Gruyter, pp. 17-54.

Johnson, M. (2010), "Cognitive science and Dewey's theory of mind, thought, and language", in M. Cochrane [Ed.], *The Cambridge Companion to Dewey*, Cambridge, Cambridge University Press, pp. 123-124.

Jung. M., Madzia, R. (2016), *Pragmatism and Embodied Cognitive Science: From Bodily Intersubjectivity to Symbolic Articulation*, Berlin, Mouton De Gruyter.

Kirchhoff, M. (2015), "Experiential Fantasies, Prediction, and Enactive Minds*", Journal of Consciousness Studies*, 22 (3-4), pp. 68-92.

Kirsh, D., Maglio, P. (1994), "On distinguishing Epistemic from Pragmatic Action", *Cognitive Science*, 18, pp. 513-549.

Kirsh, D. (2010), "Thinking with external representations", *AI and Society*, (25) 4, pp. 441-454.

Kiverstein, J., Wheeler, M. (2012) [Eds.], *Heidegger and Cognitive Science*, New York, Palgrave McMillan.

Kiverstein, J., Rietveld, E. (2012), "Dealing with context through action-oriented predictive coding. A commentary to Clark's 'Whatever the next? Predictive brains, situated agents, and the future of cognitive science'", *Frontiers in Psychology*, 3, pp. 1-2.

Koch, C., Poggio, T. (1999), "Predicting the visual world: Silence is golden", *Nature Neuroscience*, 2 (1), pp.79-87.

Lakoff, G., Johnson, M. (1980), *Metaphors we live by,* Chicago/London, The University of Chicago Press.

Lakoff, G. (1987), *Women, Fire and Other Dangerous Things. What Categories Reveal About the Mind*, Chicago/London, The University of Chicago Press.

Lakoff, G. (1990), "The Invariance Hypothesis: is abstract reason based on image-schemas?", *Cognitive Linguistics,* 1 (1), pp. 39-74.

Lakoff, G., Johnson, M. (1999) *Philosophy in the flesh. The Embodied Mind and Its Challenge to Western Thought,* New York, Basic Books.

Lakoff, G., Johnson, M. (2002), "Why cognitive linguistics requires embodied realism", *Cognitive Linguistics*, 13 (3), pp. 245-263.

Laland, K. N., Olding-Smee, J., Feldman, M. W. (2000), "Niche construction, biological evolution, and cultural change", *Behavioral and Brain Sciences*, 23, pp. 131–175.

Lashley, K. S. (1950), "In search of the engram", in F. Beach, D.O. Hebb, C. Morgan, H. Nissen [Eds.], *The Neuropshychology of Lashley*, New York, McGraw Hill, pp. 338-344.

Lazarus, R., S. (1991), *Emotion and Adaptation*, New York/Oxford, Oxford University Press.

Lauwereyns, J. (2012), *Brain and the Gaze: On the Active Boundaries of Vision,* Cambridge (MA), The MIT Press.

Lemke, J., L. (2000), "Opening Up Closure: Semiotics Across Scales", *Annals of the New York Academy of Sciences*, 901, pp. 100-111.

Lewis, C., T. (1890), *An Elementary Latin Dictionary*, New York/Cincinnati/Chicago, American Book Company.

Liszka, J. J. (1996), *A general introduction to the semeiotics of Charles Sanders Peirce*, Bloomington/Indianapolis, Indiana University Press.

Macpherson, F. (2017), "The relationship between cognitive penetration and predictive coding", *Consciousness and Cognition*, 47, p. 6-16.

Madary, M. (2015), "Extending the Explanandum for Predictive Processing. A commentary on Andy Clark", in T. Metzinger, J. M. Windt [Eds.], Frankfurt am Main: MIND Group, 7 (C), pp. 1-11, doi: 10.15502/9783958570313.

Malafouris, L. (2013), *How things shape the mind. A Theory of Material Engagement*, Cambridge (MA), The MIT Press.

Mandik, P. (2005), "Action oriented representation", in A. Brook, K. Akins [Eds.], *Cognition and the brain. The philosophy and neuroscience movement*, Cambridge (UK), Cambridge University Press, pp. 284-305.

Maran, T., Kull, K. (2014), "Ecosemiotics: main principles and current developments", *Geografiska Annaler: Series B, Human Geography*, 96 (1), pp. 41–50.

Mataric, M. (1991), "Navigating with a rat brain: A neurobiologically inspired model for robot spatial representation", in J., A, Meyer, S. Wilson [Eds.], *From Animals to Animats I,* Cambridge (MA), The MIT Press, pp. 169-175.

Maturana, H., Varela, F. (1987), *The Biological Roots of Human Understanding*, Boston, Shambhala Publications.

McDowell, J. (1994), *Mind and World*, Cambridge (MA), Harvard University Press.

Menary, R. (2006), "Attacking the bounds of cognition", *Philosophical Psychology*, 19 (3), pp. 329–344.

Menary, R. (2007), *Cognitive Integration. Mind and Cognition Unbounded,* New York, Palgrave McMillan.

Menary, R. (2009), "Intentionality, Cognitive Integration and the Continuity Thesis", *Topoi*, 28, pp. 31-43.

Menary, R. (2010a) [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press.

Menary, R. (2010b), "Cognitive Integration and the Extended Mind", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 227-244.

Menary, R. (2010c), "Introduction: The Extended Mind in Focus", in Menary, R. [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 1-28.

Menary, R. (2010d), "Introduction to the special issue on 4 E cognition", *Phenomenology and the Cognitive Sciences*, 9, pp. 459-463.

Menary, R. (2015), "What? Now. Predictive Coding and Enculturation", in T. Metzinger, J. M. Windt [Eds.], Open MIND, Frankfurt am Main: MIND Group, 25 (R), pp. 1-8, doi: 10.15502/9783958571198.

Menary, R. (2016), "Pragmatism and the Pragmatic Turn in Cognitive Science", in A. K. Engel, K. J. Friston, D. Kragic [Eds.], *The Pragmatic Turn. Toward Action-Oriented Views in Cognitive Science,* Cambridge (MA), The MIT Press, pp. 215-234.

Merleau-Ponty, M. (1942), *La structure du comportement*, Paris, PUF. / *The Structure of Behavior,* English Translation by A. Fisher, Boston, Beacon Press, 1963.

Merleau-Ponty, M. (1945), *Phénoménologie de la perception,* Paris, Gallimard. / *Phenomenology of Perception*, English translation by C. Smith, London, Routledge, 1962 (Second Edition: 2002).

Merleau-Ponty, M. (2003) *Nature: Course Notes from the Collège de France*, D. Seglard [Ed.], English translation by R Vallier, Evanston, Northwestern University Press.

Meyers, R. G. (1999), "Pragmatism and Peirce's Externalist Epistemology", *Transactions of the Charles S. Peirce Society*, 35 (4), pp. 638-653.

Miller, G. A. (2003), "The cognitive revolution: a historical perspective", *Trends in Cognitive Science*, 7 (3), pp. 141-144.

Millikan, R. (1995), "Pushmi-Pullyu Representations", *Philosophical Perspectives, 9, AI, Connectionism and Philosophical Psychology,* pp. 185-200.

Milner, D. A., Goodale, M. A. (1995), *The visual brain in action*, Oxford, Oxford University Press.

Moors, A., Ellsworth, P. C., Scherer, K. R., Frijda, N. H. (2013), "Appraisal Theories of Emotion: State of the Art and Future Development", *Emotion Review*, 5 (2), pp.119-124.

Morgagni, S. (2011), "Repenser la notion d'affordance dans ses dynamiques sémiotiques", *Intellectica*, 1 (55), pp. 241-267.

Nanay, B. (2012), "Bayes or determinables? What does the bidirectional hierarchical model of brain functions tell us about the nature of perceptual representation?", *Frontiers in Psychology*, 3, pp.1-2.

Nanay, B. (2013), *Between Perception and Action*, Oxford, Oxford University Press.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., Ric, F. (2005), "Embodiment in attitudes, social perception, and emotion", *Personality and Social Psychology Review*, 9, pp. 184-211.

Noë, A. (2004), *Action in perception*, Cambridge (MA), The MIT Press.

Noë, A. (2010), *Out of our heads. Why you are not your brain and Other Lessons from the Biology of Consciousness,* New York, Hill and Wang.

Noë, A. (2015), *Strange Tools. Art and Human Nature,* New York, Hill and Wang.

Norman, D. A. (1988), *The Design of Everyday Things*, New York, Doubleday.

Nöth, W. (1990), *Handbook of Semiotics*, Bloomington, Indiana University Press, pp. 196-197.

Nöth, W., Kull, K. (2001), "Introduction: Special issue on semiotics of nature", *Sign Systems Studies,* 29 (1), pp. 9–11.

Nöth, W. (2001), "Ecosemiotics and the semiotics of nature", *Sign Systems Studies,* 29 (1), pp. 71-81.

Nöth, W. (2016), "Habits, Habit Change, and the Habit of Habit Change According to Peirce", in D. West, M. Anderson [Eds.], *Consensus on Peirce's concept of habit. Before and beyond consciousness*, Berlin, Springer, pp. 35-63.

O' Neill, S. (2008), *Interactive Media. The Semiotics of Embodied Interaction*, London, Springer.

O'Regan, K., Noë, A. (2001), "A sensorimotor approach to vision and visual consciousness", *Behavioral and Brain Sciences,* 24 (5), pp. 883–975.

Orlandi, N. (2016), "Bayesian perception is ecological perception", *Philosophical Topics*, 44 (2), pp. 327-351.

Paolucci, C. (2008), "From Logic of Relatives to Cognitive Semiotics: On some Unsuspected Correspondences between Peirce and Structuralism", *Cognitive Semiotics*, 3, pp. 91-113.

Paolucci, C. (2010), *Strutturalismo e Interpretazione*, Milano, Bompiani.

Paolucci, C. (2011), "The External Mind: Semiotics, Pragmatism, Extended Mind and Distributed Cognition", *Versus-Quaderno di Studi Semiotici*, 112-113, pp. 69-96.

Pape, H. (2014), "Peirce's Process Ontology of Relational Order", in T. Thellefsen, B. Sørensenm [Eds.], *Peirce in his own words. 100 years of semiotics, communication, and cognition*, Berlin, Mouton de Gruyter, pp. 239-244.

Parker, K. (1994), "Peirce's Semeiotic and Ontology", *Transactions of the Charles S. Peirce Society*, 30 (1), pp. 51-75.

Pattee, H. (1995), "Evolving Self-Reference: Matter, Symbols and Semantic Closure", *Communication and Cognition – Artificial Intelligence,* (12), pp. 9 – 27.

Peirce, C., S. (1931-58), *Collected Papers of Charles S. Peirce*, C. Hartshorne, P. Weiss, A. W. Burks [Eds.], Cambridge (MA), Harvard University Press.

Peirce, C., S. (1984), *Writings of Charles Sanders Peirce. A Chronological Edition,* Volume II, 1867-1871, C. Moore [Ed.], Bloomington, Indiana University Press.

Peirce, C., S. (1986), *Writings of Charles Sanders Peirce. A Chronological Edition,* Volume III, 1872-1878, C. J. W. Kloesel [Ed.], Bloomington, Indiana University Press.

Peirce, C., S. (1989), *Writings of Charles Sanders Peirce. A Chronological Edition,* Volume IV, 1879-1984, J. W. Kloesel [Ed.], Bloomington, Indiana University Press.

Peirce, C., S. (1993), *Writings of Charles Sanders Peirce. A Chronological Edition,* Volume V, 1884-1886, J. W. Kloesel [Ed.], Bloomington, Indiana University Press.

Peirce, C., S. (1998), *The Essential Peirce, Volume 2: Selected Philosophical Writings,* 1893-1913, Edited by the Peirce Edition Project, Bloomington, Indiana University Press.

Peirce, C., S. (2010), *Writings of Charles Sanders Peirce. A Chronological Edition,* Volume VIII, 1890-1892, Edited by the Peirce Edition Project, Bloomington, Indiana University Press.

Petitot, J. (1985), *Morphogenèse du sense* I, Paris, PUF.

Prinz, J. (2003), "Emotions Embodied", in R. Solomon [Ed.], *Thinking about Feeling*, New York, OUP, pp.44-61.

Prinz, J. (2004), *Gut reactions. A perceptual theory of emotion,* Oxford, Oxford University Press.

Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R. J. (2005a), "Brain signatures of meaning access in action word recognition", *Journal of Cognitive Neuroscience,* 17 (6), pp. 884-892.

Putnam, H. (1973), "Meaning and Reference", *Journal of Philosophy*, 70, pp. 699–711.

Queiroz, J., Merrell, F. (2008), "On Peirce's Pragmatic Notion of Semiosis. A contribution for the design of meaning machines", *Minds and Machines*, 9 (1), pp. 129-143.

Ramsey, W. M. (2007), *Representation reconsidered*, Cambridge, Cambridge University Press.

Rao, R., P., Ballard, D., H. (1998), "Predictive coding in the visual cortex: a functional interpretation of extra-classical receptive fields effects", *Nature Neuroscience*, 2, pp. 79-87.

Ratcliffe, M. (2002), "Heidegger's attunement and the neuropsychology of emotion", *Phenomenology and the Cognitive Sciences*, 1, pp. 287–312.

Ratcliffe, M. (2013), "Why Mood Matters", in M. Wrathall [Ed.], *The Cambridge Companion to Heidegger's Being and Time*, Cambridge, Cambridge University Press, pp. 157-176.

Reed, E. S. (1996), *Encountering the world: Toward an ecological psychology*, Oxford, Oxford University Press.

Rietveld, E. (2008a), "The skillful body as a concernful system of possible actions: Phenomena and neurodynamics", *Theory & Psychology*, 18 (3), pp. 341-63.

Rietveld, E. (2008b), "Situated normativity: the normative aspect of embodied cognition in unreflective action", *Mind*, 117 (468), pp. 973-1001.

Rietveld, E. (2012a), "Context-switching and responsiveness to real relevance", in J. Kiverstein, M. Wheeler [Eds.], *Heidegger and Cognitive Science: New Directions in Cognitive Science and Philosophy*, Basingtoke, Hampshire, Palgrave Macmillan, pp. 105-135.

Rietveld, E., Kiverstein, J. (2014), "A rich landscape of affordances", *Ecological Psychology*, 26 (4), pp. 325-352.

Robbins, P., Aydede, M. [Eds.] (2009), *The Cambridge Handbook of Situated Cognition*, Cambridge (MA), Cambridge University Press.

Robin, R. S. (1967), *Annotated Catalogue of the Papers of Charles S. Peirce,* Cambridge, The University of Massachusetts Press.

Roll, J-P., Roll, R. (1988), "From eye to foot: A proprioceptive chain involved in postural control", in G. Amblard, A. Berthoz, and F. Clarac [Eds.], *Posture and Gait: Development, Adaptation, and Modulation,* Amsterdam, Excerpta Medica, pp. 155-164.

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

Rooij, I., Bongers, R. M., Haselager, W. F.G. (2002), "A non-representational approach to imagined action", *Cognitive Science*, 26, pp. 545-375.

Roy, J. M. (2013), "Cognitive Pragmatism and Enactivism", *Intellectica*, 60, pp. 69-90.

Rowlands, M. (2006a), *Body language. Representation in action*, Cambridge (MA), The MIT Press.

Rowlands, M. (2006b), "The normativity of action", *Philosophical Psychology*, 19 (3), pp. 401-416.

Rowlands, M. (2009), "Situated representations", in P. Robbins, M. Aydede, M. [Eds.], *Cambridge Handbook of Situated Cognition*, Cambridge, Cambridge University Press, pp.117-133.

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

Rowlands, M. (2012), "Representing without representations", *Avant: Trends in Interdisciplinary Studies*, 3 (1), pp. 133-144.

Rowlands, M. (2015), "Arguing about representations", *Synthese*, Online First, pp.1-18.

Rupert, R. D. (2004), "Challenges to the hypothesis of extended cognition", *Journal of Philosophy,* 101 (8), pp. 389-428.

Rupert, R. D. (2009), *Cognitive Systems and the Extended Mind*, Oxford, Oxford University Press.

Salgado-Pineda, P., Delaveau, P., Blin, O., Nieoullon, A. (2005), "Dopaminergic Contribution to the Regulation of Emotional Perception", *Clinical Neuropharmacology,* 28, pp. 228–237.

Santaella, L. (2015), "Sound and Music in the Domain of Rhematic Iconic Qualisigns", *Signata*, 6, pp. 91-106.

Schmitt, R. (1965), "Heidegger's Analysis of 'Tool'", *The Monist*, 49 (1), *Linguistic Analysis and Phenomenology*, pp. 70-86.

Searle, J. (1980), "Minds, Brains and Programs", *Behavioral and Brain Sciences*, 3, pp. 417–457.

Sellars, W. (1997), *Empiricism and the Philosophy of Mind*, Cambridge (MA), Harvard University Press.

Shapiro, L. A. (2004), *The Mind Incarnate*, Cambridge (MA), The MIT Press.

Shapiro, L. A. (2007), "The Embodied Cognition Research Programme", *Philosophy Compass*, 2 (2), pp. 338-346.

Shapiro, L. A. (2011), *Embodied Cognition*, New York, Routledge.

Shapiro, L. A. (2014) [Ed.], *The Routledge Handbook of Embodied Cognition*, London/New York, Routledge.

Shook, J. R., Solymosi, T. (2014a) [Eds.], *Neuroscience, Neurophilosophy and Pragmatism. Brains at work with the world*, New York, Palgrave Macmillan.

Shook, J. R., Solymosi, T. (2014b) [Eds.], *Pragmatist Neurophilosophy. American Philosophy and the Brain*, New York, Bloomsbury.

Short, T., L. (1996), "Interpreting Peirce's Interpretant", *Transactions of the Charles S. Peirce Society*, 32 (4), pp. 488-541.

Short, T., L. (2007), *Peirce's Theory of Signs,* Cambridge, Cambridge University Press.

Silva, P., Garganta, J., Araújo, D., Davids,K., Agular, P. (2013), "Shared knowledge or shared affordances? Insights from an ecological dynamics approach to team coordination in sports", *Sports Medicine*, 34, pp. 765-772.

Simon, H. (1956), "Rational choice and the structure of the environment", *Psychological Review*, 63 (2), pp. 129-138.

Skagested, P. (2004), "Peirce's Semeiotic Model of the Mind", in C. Misak [Ed.], *The Cambridge Companion to Peirce,* Cambridge (MA), Cambridge University Press, pp. 241-256.

Smith, C. A., Lazarus, R. S. (1993), "Appraisal Components, Core Relational Themes, and the Emotions", *Cognition and Emotions*, 7 (3/4), pp. 233-269.

Sonesson, G. (2012), "The foundation of cognitive semiotics in the phenomenology of signs and meanings", *Intellectica*, 58 (2), pp. 207-239.

Spratling, M. W. (2015), "A review of predictive coding algorithms", *Brain and Cognition*, 112, pp. 92-97.

Stapleton, M., Ward, D. (2012), "Es are good. Cognition as enacted, embodied, embedded, affective and extended", in F. Paglieri [Ed.], *Consciousness in Interaction. The role of the natural and social context in shaping consciousness*, Amsterdam, John Benjamins Publishing Company, pp. 89-104.

Steiner, P. (2008), "Sciences cognitives, tournant pragmatique et horizons pragmatistes", *Tracés. Revue de Sciences Humaines*, 15, pp. 85-104.

Steiner, P. (2013), "Pragmatism(s) and Cognitive Science: Introductory Remarks", *Intellectica*, 60, pp. 7-48.

Sterelny, K. (2010), "Minds: Extended or Scaffolded?", *Phenomenology and the Cognitive Sciences*, 9, pp. 465-481.

Stutz, A. J. (2014), "Embodied niche construction in the hominin lineage: semiotic structure and sustained attention in human embodied cognition", *Frontiers in Psychology*, 5, pp. 1-19.

Sutton, J. (2010), "Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process," in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 189-225.

Sutton, J. (2014), "Remembering as public practice: Wittgenstein, memory and distributed cognitive ecologies", in D. Moyal-Sharrock, V. A. Munz, A. Coliva [Eds.], *Mind, Language, and Action: Proceedings of the 36th Wittgenstein Symposium*, Berlin, Mouton de Gruyter, pp. 409-44.

Taylor, J. R. (2007), "Cognitive Linguistics and Autonomous Linguistics", in Geeraerts, D., Cuyckens, H. [Eds.], *The Oxford Handbook of Cognitive Linguistics*, Oxford, Oxford University Press, pp. 566-587.

Thompson, E. (2004), "Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela", *Phenomenology and the Cognitive Sciences*, 3, pp. 381- 398.

Thompson, E. (2007), *Mind in Life. Biology, Phenomenology, and the Sciences of Mind,* Cambridge (MA), Harvard University Press.

Thompson, E., Stapleton, M. (2009), "Making sense of sense-making: Reflections on enactive and extended mind theories", *Topoi,* 28, pp. 23-30.

Thompson, E., Cosmelli, D. (2011), "Brain in a Vat or Body in a World? Brainbound versus Enactive Views of Experience", *Philosophical Topics*, 39 (1), pp. 163-180.

Tiercelin, C. (1995), "The relevance of Peirce's Semiotic for Contemporary Issues in Cognitive Science", *Acta Philosophica Fennica. Mind and cognition: philosophical perspectives on cognitive sciences and artificial intelligence*, 58, pp. 37-73.

Tribus, M. (1961), *Thermodynamics and thermostatic: An introduction to energy, information and states of matter, with engineering applications*, New York, D. Van Nostrand.

von Uexküll, J.V. (1984), "The Theory of Meaning", *Semiotica*, 42 (1), pp. 25-82.

Varela, F. (1984), "Living ways of sense-making: a Middle Path for Neurosciences", in P. Livingstone [Ed.], *Disorder and Order,* ANMA, Saratoga, pp. 41-60.

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

Varela, F. (1999), "The specious present : a neurophenomenology of time consciousness", in J. Petitot, F. Varela, B. Pachoud, J.M. Roy [Eds.], *Naturalizing phenomenology. Issues in contemporary phenomenology and cognitive science*, Stanford, Stanford University Press, pp. 266-329.

Violi, P. (2003), "Le tematiche del corporeo nella linguistica cognitive", in L. Gaeta, S. Luraghi [Eds.], *Introduzione alla linguistica cognitiva*, Roma, Carocci, pp. 57-76.

Violi, P. (2008), "Beyond the body: Towards a full embodied semiosis", in F. R. Dirven [Ed.], *Body, language and mind*, vol. 2, Berlin, Mouton de Gruyter, pp. 241-264.

Walter, S. (2010), "Cognitive extension: the parity argument, functionalism, and the mark of the cognitive", *Synthese*, 177, pp. 285–300.

Ward, D., Roberts, T., Clark, A. (2011), "Knowing what we can do: actions, and the construction of phenomenal experience", *Synthese*, 181 (3), pp. 375-394.

Ward, D., Silverman, D., Villalobos, M. (2017), "Introduction: The Varieties of Enactivism", *Topoi*, pp.1-11.

West, D. (2016), "Indexical Scaffolds to Habit Formation", in D. West, M. Anderson [Eds.], *Consensus on Peirce's concept of habit. Before and beyond consciousness*, Berlin, Springer, pp. 215-240.

Wheeler, M. (2005), *Reconstructing the cognitive world. The next step,* Cambridge (MA), The MIT Press.

Wheeler, M. (2008a), "Minimal representing: A response to Gallagher", *International Journal of Philosophical Studies*, 16 (3), pp. 371-376.

Wheeler, M. (2008b), "Cognition in context: phenomenology, situated robotics and the frame problem", *International Journal of Philosophical Studies*, 16 (3), pp. 323-349.

Wheeler, M. (2010a), "The problem of representation", in S. Gallagher S., D. Schmicking [Eds.], *Handbook of Phenomenology and Cognitive Science*, Berlin, Springer, pp. 319-36.

Wheeler, M. (2010b), "In defense of extended functionalism", in R. Menary [Ed.], *The Extended Mind*, Cambridge (MA), The MIT Press, pp. 245-270.

Wickelgren, I. (1997), "Getting the brain's attention", *Science*, pp. 35-37.

Williams, D. (2017), "Predictive Processing and the Representation Wars", *Minds & Machines*, Online First, pp. 1-32.

Wilson, R. A. (1994), "Wide computationalism", *Mind*, 103 (411), pp. 351-372.

Wilson, M. (2002), "Six views of embodied cognition," *Psychonomic Bulletin and Review*, 9, pp. 625–636.

Windsor, W. L. (2004), "An Ecological Approach to Semiotics", *Journal for the Theory of Social Behavior*, 34 (2), pp. 179-198.

Winston P. H. (1984), *Artificial Intelligence,* Second Edition, Reading (MA), Addison-Wesley.

Wrathall, M., Malpas, J. (2000), *Heidegger, Coping and Cognitive Science*, Cambridge (MA), The MIT Press.

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

Zlatev, J. (2012), "Cognitive semiotics: an emerging field for the transdisciplinary study of meaning", *The Public Journal of Semiotics*, 4 (1), pp. 2-24.

Web sources

Commens. Digital Companion to C. S. Peirce (http://www.commens.org/encyclopedia/article/fabbrichesi-rossella-peirce-mead-and-theory-extended-mind)

The Concise Oxford Dictionary of English Etymology (http://www.oxfordreference.com/view/10.1093/acref/9780192830982.001.0001/acref-9780192830982)

The Online Etymology Dictionary (http://www.etymonline.com/index.php)

The Oxford English Dictionary (http://www.oed.com/)