



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

Pierre Steiner

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

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

Abstract

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

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

Introduction

The criticism and the rejection of representationalism have a particular status in enactivism (or enaction). The definition of cognition as *embodied action* was explicitly proposed by Varela, Thompson and Rosch in their seminal *The Embodied Mind* as an alternative to the definition of cognition as “the representation of a world that is independent of our perceptual and cognitive capacities by a cognitive system that exists independent of the world” (1991: xx). Still, if we are looking nowadays for a more positive definition of enaction, concepts such as *autonomy*, *autopoiesis*, *embodiment*, *structural coupling*, *sense-making*, *life-mind continuity*, or *lived experience* immediately and eminently come to the fore, somehow leaving the criticism of representationalism in the back-

ground¹⁷. It is true that the unpacking of the meanings of these concepts often carries or implies an implicit criticism of representationalism as the one mentioned above, and is very helpful for contrasting enactivism with other non-representationalist theories of cognition such as radical embodied cognitive science (Chemero 2009). Nevertheless, enactivists would be wrong to consider that developing an explicit criticism of representationalism was only necessary when enaction emerged as an alternative theory (or even paradigm) to cognitivism, where the concept of “mental representation” was—and is still—a basic building block. Strategically, the rejection of representationalism continues to mark an important difference not only between enactivism and cognitivism (i.e. the computo-representational theory of mind), but also between (some forms of) enactivism and other more recent theories that criticize, amend or even reject the intellectualist, internalist or formalist dimensions of the computo-representational theory of mind... by often retaining representationalism. Amongst these recent theories, one can include distributed cognition (Hutchins 1995), situated cognition (Clancey 1997), extended cognition (Clark 2008), and interactivism (Bickhard 2009)¹⁸. More fundamentally, the permanent rejection of representationalism should not be taken as an *easy* task: representationalism is a polymorphous and plastic thesis, sometimes looking like a Lernaean Hydra: shallow and maximalist versions of representationalism are easy to see and to dislodge, but only at the benefit of other versions which are much more tenacious, refined and pervasive. It is the implicit endorsement of these latter versions that may explain why some authors propose “enactive” accounts of mental representations (Ellis and Newton 2010), or that some reviewers of the enactivist literature state that “nothing in the enactivist view requires abandonment of contentful states” (Shapiro 2014).

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

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

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

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

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

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

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

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

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

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

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

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

Table 1

By looking at this chart, one may note that radical enactivism is rejecting a very basic kind of representationalism which is logically narrower than the representationalism just defined above, but which also constitutes the core assumption of the great majority of actual versions of representationalism: the existential claim that there are mental representations, defined as physical structures (vehicles) playing a role in cognitive processing *in virtue of some content whose existence does not depend on the existence of social and linguistic practices and on the ability of the organism to take part in those practices*¹⁹. Typically, these mental representations consist in subpersonal and intracranial processes that are naturally or intrinsically contentful²⁰. Radical enactivism refuses that idea by arguing that

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

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

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

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

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

Table 2

In this paper, I wish to clarify and to assess some arguments proponents of enactivism (radical or not) have proposed in their criticism of representationalism (in the narrow sense just defined above). This will first require a set of distinctions about the *targets* and the *forms* of enactive anti-representationalism (section I). I will then proceed by rehearsing two *classical* strategies

against representationalism that the enactivist tradition has exploited but that can also be found in other traditions (section II), before presenting two more *radical* and *specific* anti-representationalist strategies that I see as being more proper to the enactivist tradition (section III). Still, these two latter strategies might pose no problems for a marginal yet existing version of representationalism, according to which mental representations do not essentially have content and intentionality (section IV). This resistance of a marginal version of representationalism to the (proclaimed) radicality of enactive anti-representationalism will absolutely not be considered here as a refutation or dismissal of the latter. On the contrary, it will be seen as an opportunity—in section V—to identify some common assumption(s) that enactive anti-representationalism and classical representationalism might share, but also to invite enactivist anti-representationalists to (re)consider the conditions by which theoretical terms may be eliminated (or retained) in science.

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

I. Situating anti-representationalism

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

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

A very basic distinction is classically drawn between *methodological* and *ontological* versions of representationalism and anti-representationalism²². Ontological versions explicitly take issue with the reality of mental representations. Ontological representationalism considers that mental representations exist, whereas ontological anti-representationalism denies their existence. A methodological version is agnostic on the issue of the reality of mental representations: they are not taken as literally existing (we may not be warranted in positively asserting their reality). Still, methodological representationalism argues that mental representations (and their manipulation, manufacture or retrieval) must necessarily be posited for the explanation or prediction of the performances of cognitive systems; whereas methodological non-representationalism argues that they do not need to be invoked (it is possible to posit them, but there are more helpful theoretical posits). In both methodological stances, nothing ontological is inferred from the presence or the absence of the concept “mental representation” in successful explanatory and predictive practices (methodological representationalism includes “fictionalism” about mental representations (Sprevak, 2013)). Methodological versions of representationalism and anti-representationalism may be said to be anti-realist in the following sense: they deny that theories involving the positing (or the non-positing) of mental representations are truth-conditioned descriptions of their intended domain (observable and unobservable), and that their predictive or explanatory successes entail that the entities they posit have “real” counterparts.

Methodological *non*-representationalism is not a variety of *anti*-representationalism, since it basically makes *no* use of the concept “mental representation”. In itself, it is not *against* the existence of mental representations: it may consider that the ontological debate between representationalism and anti-representationalism is vain, for instance because of a lack of clear definition of what a representational property is²³. Anti-representationalism is more demanding and challenging than non-representationalism, since it explicitly claims that mental representations do not exist.

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

²² See Chemero (2000) for this distinction (but I use the term “methodological” where Chemero used “epistemological”).

²³ See Haselager et alii. (2003).

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

We thus have eight possible positions:

- Ontological representationalism;
- Methodological representationalism (which I will denote hereafter by “methodological representationalism”);
- Ontological representationalism with methodological representationalism (which I will denote hereafter by “representationalism”);
- Ontological anti-representationalism;
- Methodological non-representationalism (which I will denote hereafter by “methodological non-representationalism”);
- Ontological anti-representationalism with methodological non-representationalism (which I will denote hereafter by “anti-representationalism”);
- Ontological representationalism with methodological non-representationalism: one accepts that mental representations exist, but consider that scientific models should better do without the concept of “mental representation”.
- Ontological anti-representationalism with methodological representationalism: one accepts that mental representations do not exist, but still holds that they are our best ways to capture and explain the complexity of cognitive behaviour.

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

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

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

We thus have twenty-four (8 x 3) different positions. I will focus on what I consider as being the twelve most notable positions. Indeed, for obvious reasons of space but also of relevance, I leave aside global, basic and local versions of ontological representationalism and anti-representationalism: in the

philosophy of cognitive science, there are not many scholars who defend ontological representationalism or anti-representationalism *only*, not aggregating them with methodological commitments or suggestions. Ontological representationalism and anti-representationalism are almost always included as components of what I call here “representationalism” and “anti-representationalism”, which also include methodological commitments. I also leave aside global, basic and local versions of conjunctions of ontological representationalism with methodological non-representationalism, and of ontological anti-representationalism with methodological representationalism. These versions are quite rare in the literature, and it is hoped that the reader will be able to define them from the statements given above. This leaves us with 12 positions, which we will now examine one-by-one.

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

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

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

All these cases of representationalism are composed of ontological representationalism with methodological representationalism. Mental representations are here contentful (information-carrying) physical structures that have a real ontological and explanatory status. Their content is a natural product that allows them to refer to some object, property or state of affairs. According to this definition, external (public and/or shareable) representations such as models, images or natural language sentences are *not* mental representations: one can be a critic of representationalism without denying the existence and the cognitive importance of external representations. Representationalism is

here a claim about the *reality* of the property “being a mental representation of X”: it is not only about its heuristic relevance for describing and explaining the mechanisms of cognition. Representationalism, as we define it here, is a very general claim, that includes many variations: one can find a version of representationalism for which all mental representations are necessarily symbols in a language of thought (Fodor, 1987, chap.1 and appendix), or a version of representationalism for which cognition must be defined as a set of operations having the function of building mental representations of environmental phenomena (in that version, the property of mental representation is used for defining the *explanandum*, and not only the *explanantia* of cognitive science). These two versions of representationalism can be criticized, amended or even rejected by other versions of representationalism (for instance: parallel and sub-symbolic distributed processing vs. the symbolic conception; action-oriented conceptions of cognition and representation vs. the idea of cognition as a mirror of the environment ; mental representations as maps, models or pictures vs. mental representations as propositional sets of symbols...). Global representationalism does not necessarily link the cognitive character of a system to the presence of mental representations: it just asserts that from some level of study and analysis of cognitive systems, it is necessary to acknowledge the existence of mental representations, without assuming that they correspond perfectly to what would be described at another level of analysis of these same systems. Still, global representationalism considers that explaining the cognitive properties of cognitive systems requires the appeal to mental representations.

Vehicle-internalist and cognitivist theories of cognition are not the exclusive owners of representationalism. On the contrary: many versions of extended, distributed or situated cognition may endorse basic representationalism. For most proponents of extended cognition, even if there may be cases or aspects of extended cognitive processing that do not (just) involve mental representations as made out of natural content and thus that do not require representationalist explanations, the existence of mental representations (for instance realized in intracranial and subpersonal processes) and the necessity of referring to them when one tries to explain a great variety of cognitive phenomena are not at issue²⁴. Unlike classical AI representations, these representations do not need to be complete, inert, propositional, denotational, action- and perception-neutral, stable, complex, detailed, digital, discrete, amodal, syntactically structured, or symbolic. Mental representations can be built and used on the fly; they can be modal (even when they are categorical), minimal (content-sparse), superposed, partial, action-oriented, context-dependent, embodied, analogue, distributed, or sub-symbolic. Basic representationalism is more

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

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

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

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

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

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

I mentioned above that methodological versions of representationalism (and anti-representationalism) were anti-realist towards the property “being a mental representation”. There are different kinds of anti-realism: instrumentalism, but also pragmatism, phenomenism, interpretationism, constructive empiricism, fictionalism and idealism, to name but a few. The combination of methodological representationalism with each of these versions of anti-realism can foster complex and subtle versions of methodological representa-

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

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

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

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

The three cases of non-representationalism are not symmetrical with the three cases of methodological representationalism, since methodological representationalisms express a necessity in the form of a normative claim (“one *must* appeal to the concept of ‘mental representation’”) while non-representationalisms deny this necessity without implying that non-representationalist models and explanations should automatically be *preferred* to representationalist models and explanations. And remember that

non-representationalism do not claim that representationalist explanations are *mistaken*. It just holds that they can be dispensed with (globally, basically or locally).

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

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

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

Since representationalism and anti-representationalism result from a conjunction of ontological and methodological positions, each of these latter positions can *partially* support representationalism and anti-representationalism. Global representationalism and global anti-representationalism are more than methodological positions, and concern every cognitive system: empirical statements and examples will not be sufficient for justifying them. Moreover, global, basic and local versions of anti-representationalism include ontological claims on the non-reality of mental representations. These ontological claims will be conceptual, not empirical, for one cannot ask to the proponent of anti-representationalism to empirically show that mental representations do not exist²⁵. But anti-representationalisms also include methodological components: arguments related to the possibility of non-representational explanatory practices in cognitive science can partially justify them.

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

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

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

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

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

(A) Among the six different versions of representationalism (1, 2, 3, 4, 5, and 6), which version(s) does enactivism reject?

(B) Among the six different versions of anti-representationalism (7, 8, 9, 10, 11, and 12), which version(s) does enactivism endorse?

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

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

The first trend is an explicitly *eliminativist* trend, consisting in the explicit defence of global anti-representationalism, and thus in the complete rejection of all the six versions of representationalism, ontological and methodological. Hutto and Myin's *radical enactivism* is here a landmark. The battle against representationalism is here engaged on two fronts: an ontological front (mental representations as made out of natural content do not exist), and a methodological front (non-representationalist explanations are possible and should

be favored). But we must keep in mind that radical enactivism is *not* “really radical enactivism”: radical enactivism accepts that some classes of cognitive phenomena exhibit representational properties, but these representational properties necessarily depend on the existence of sociocultural practices (representational properties are not natural properties). Since radical enactivism denies the very existence of natural content, it is coherent here to define it as defending a global version of anti-representationalism (as we have defined it).

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

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

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

The conservative trend faces at least two dangers: first, the proponents of basic representationalism (be it ontological and methodological or methodo-

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

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

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

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

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

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

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

The existence of these three trends *within* enactivism is both normal and problematic. It is normal, for it would be naive (and even dangerous) to expect too much homogeneity from enactivism. There are different ways of embracing the computational theory of mind or distributed cognition: why would that be different for enactivism? But it is also problematic, for these

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

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

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

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

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

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

II. Philosophical and explanatory strategies against representationalism

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

Dreyfus, Charles Taylor, Hilary Putnam, Robert Brandom, John McDowell or Charles Travis have developed non-representationalist approaches of conceptual, perceptual, doxastic or epistemic intentionality. The philosophical strategy that enactivism can deploy (and has deployed) against representationalism often relies on the works of some of these authors (see for instance Varela, Thompson and Rosch 1991). Nevertheless, this strategy is not without certain defects. The main potential problem seems to be the following: this strategy is based on philosophical considerations about intentionality, cognition, knowledge or reasoning as *personal-level* phenomena. These considerations do not seem to entail a rejection of representationalism as a definition of what makes these phenomena causally possible on a subpersonal level²⁶. On the contrary: representationalism, as a hypothesis concerning the existence of subpersonal (and possibly intracranial) mental representations endowed with natural content, may always be invoked for explaining for instance *how* personal-level perception, cognition or intentionality is for action or develops from practical coping or embodiment. This philosophical strategy may also be closer to *really radical enactivism* than to *radical enactivism*: the works of the philosophers mentioned above may include criticisms of the idea that non-basic forms of cognition involve *representational* content, even if the latter one is defined from social, linguistic or cultural resources.

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

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

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

sentations at all (and not only explicit representations, as in Brooks' famous case), but also the study of insect cognition (Webb, 1994), developmental processes (Thelen & Smith 1994), motor coordination, or perceptual supplementation (Lenay & Steiner 2010). If there should be an empirically robust basic methodological non-representationalism, it will be built out of *many* cases of local methodological representationalism²⁷. But this hypothetical inductive generalization is not the only challenge that is encountered by this explanatory strategy against representationalism. Indeed, according to its representationalist opponents, it must also show that

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

and that

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

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

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

The explanatory strategy can also take the form of an original criticism of representationalism made by William Ramsey (2007): for Ramsey, methodological non-representationalism is already implicitly dominant in cognitive science, in spite of the representationalist propaganda. Indeed, if one pays attention to the properties in virtue of which the entities named “mental representations” play an explanatory role in many models of cognitive behaviour, one will see how much these properties are very rarely representational properties (even if the observer can ascribe them semantic properties): these properties are more elemental, since they often only take the form of indication, covariation, correlation or standing-in. If Ramsey’s arguments are correct, many proclaimed “representationalist” explanations of cognitive phenomena are actually non-representational, and form a set of examples that can be exploited by the proponent of explanatory anti-representationalism²⁸. Following Ramsey or Gallagher (2008), one can for instance wonder how much the so-called “minimal representations” that are defended by proponents of extended cognition like Mike Wheeler or Mark Rowlands are robust enough for having representational properties.

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

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

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

III. Ontological and epistemological strategies against representationalism

These two strategies support the idea that representational content²⁹ does not exist at a natural and subpersonal level, so that—by implication—(naturally) *contentful* physical structures do not exist. These two strategies are sufficient for excluding ontological representationalism, but are not sufficient for ruling out methodological versions of representationalism: only methodological non-representationalism (with the help of the explanatory strategy) can do that. In the enactive literature, these two strategies are included in global versions of anti-representationalism, ontological and methodological. But since these two strategies (only) concern the existence of mental representations, they could also be embedded in positions such as ontological anti-representationalism only, or ontological anti-representationalism *with* methodological representationalism (see section II).

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

Let us begin by the epistemological strategy.

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

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

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

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

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

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

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

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

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

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

Maturana already expressed this epistemological strategy in 1978:

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

This argument is close to, but not identical with, a basic anti-representationalist argument in the Wittgensteinian tradition (see for instance Kenny 1989, chap. 10; Glock 2008; Descombes 2010; Bennett and Hacker 2003). According to this argument, it is a category mistake to think that natural and subpersonal phenomena of covariation or causal dependency could harbour or amount to representational or semantic properties. The latter properties are necessarily derived from linguistic and social practices, in which shareable structures (pictures, sentences, models,...) can acquire a representational status in virtue of what rule-following agents do and must do with them. The existence of mental representations as made out of natural content is here *a priori* excluded from the stipulation of necessary conditions for the existence

of semantic content. These conditions will never be satisfied by a brain (even if having a brain is a necessary causal condition for being able to participate to linguistic and social practices from which representational properties emerge). Both sides (Maturana-Varela *and* the Wittgensteinian side) agree that semantic content is a product of interpersonal and normative practices. A (inter-)personal-level entity cannot be used as the general blueprint for explanations or descriptions at the subpersonal level. The Wittgensteinian argument focuses on what brains *cannot do* so that it is fallacious to ascribe representational properties to them, while Varela and Maturana focus more on what the observer *is doing* when he is using the representationalist talk for describing the brain, in order to show how representationalism rests on an illusion.

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

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

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

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

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

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

P1. In the representationalist ontology, the subpersonal and intracranial phenomena named “mental representations” naturally (or intrinsically) have contents³¹

P2. There is no natural (or intrinsic) content at the level of subpersonal and intracranial phenomena

C. Subpersonal and intracranial mental representations, as they are conceived by the representationalist ontology, do not exist

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

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

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

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

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

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

What are the conditions in virtue of which one can say that some entity or property does not exist, rather than say that it exists, although it is very different from what one thought and thinks about it? What are the conditions in virtue of which some term does not refer to anything, rather than referring to something which is very different from what the descriptions associated with it prescribe? There is no definite answer(s) to these questions: a consideration of the properties of the theory in which the term is defined may be crucial, but these properties can be and have been very different from case to case. The

result (that is, the elimination or the conservation of the term) can derive from the fact the theory is seen as true or false, too simple or too complicated, necessary (being the only game in town) or useless, reducible or not to a new theory... but it can also be generated by social and pragmatic factors. It is very easy to assert, *ex post facto*, that eliminated scientific concepts *had to* be eliminated because right from the start, nothing corresponded to them. But at the time of the controversy between their proponents and their opponents, this alleged absence of reference was exactly the disputed issue: it was a disputed argument, not an undisputed conclusion. What firstly motivated the effective elimination of (pseudo)scientific concepts was generally not the inexistence of their referents (how could something inexistent make by itself a causal difference?), and not only the availability and the relevance of other concepts (“oxygen” over “phlogiston” for instance), for the concepts that finally became eliminated were available and theoretically relevant: it was also a set of interests, purposes and institutional factors which increasingly led to their elimination. A classical proponent of eliminativism, Patricia Churchland, was clearly aware of the heterogeneity of the causes that can contribute—or not—to the elimination of a theory or of a theoretical term when she wrote that

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

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

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

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

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

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

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

1. Mental representations are distally interpreted: they are about objects and properties in the environment.
2. Computational states and processes are type-individuated with reference to these distal objects and properties³²
3. The relation between mental representations and the distal objects and properties to which they are mapped is a substantive, naturalistically specifiable relation.

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

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

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

Egan's proposal is both progressive and conservative: progressive, in so far as she argues that distal content does not have causal and individuating roles in computational models of cognitive processing. For this reason, her version of representationalism avoids most of the criticisms of the ontological and of the epistemological strategies: the relation between mental representations and distal objects is not a real and naturalistically specifiable relation, since distal content is ascribed to these mental representations. This ascription of distal content is based on the fact these internal states and structures co-vary with environmental properties and facts; but these states and structures do not represent their normal distal cause (Egan, 2010: 257). Still, Egan's position is conservative with respect to the notion of content: even *unreal* (that is, not grounded on a natural relation, and not having a causal and individuating role) and defined as the product of an ascription, content must be retained³³. It is the ascription of content which defines what a *cognitive* mechanism or task is. (The ascription of) content is therefore methodologically necessary. Mental representations exist (they are not ascribed), although their content is (only) ascribed: the explanatory strategy can criticize the latter claim, while the ontological strategies seen above cannot attack the first claim.

Chomsky is much more radical than Egan, by defending the idea that content *tout court*, including the concept of "content", has no place at all in cognitive science³⁴. Indeed, for a long time now, Chomsky is convinced that folk and philosophical notions like "content", "intentionality" and "reference" have no place at all in a naturalistic framework dedicated to the understanding of cognitive faculties. There is thus no need to discuss the current failures of attempts to naturalize representational content for dropping out content in the definition of mental representations:

The central problem that troubles me is this. I do not know of any notion of 'representational content' that is clear enough to be invoked in accounts of how internal computational systems enter into the life of the organism. And to the extent that I can grasp what is intended, it seems to be very questionable that it points to a profitable path to pursue. (Chomsky 2003a: 274).

To be an Intentional Realist, it would seem, is about as reasonable as being a Desk- or Sound-of-Language- or Cat- or Matter-Realist; not that there are no such things as desks, etc., but that in the domain where questions of realism arise in a serious way, in the context of the search for laws of nature, objects are not conceived from the peculiar perspectives provided by concepts of common-sense. (...) Intentional phenomena relate to people and what they do as viewed

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

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

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

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

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

“Informal usage”, here, means the very widespread tendency to embrace a linguistic model of mental representations, assuming they have semantic content or truth conditions, like daily linguistic products. The proponent of contentless representationalism, here, wants to sever the ties between the explanatory posits of cognitive science and our folk understanding of what representations are, but considers that the very term of “representation” can

still be applied for describing the natural, intracranial and subpersonal reality he is studying (even though its meaning, in scientific inquiry, is very different from our daily conceptions). Contentless representations entertain functional relations with external phenomena: they occur when and only when the organism interacts or deals with these external phenomena. In this sense, a “number-representation” is a representation of a different functional type than a “face-representation”, but is not to be defined as a representation of an external item. But—and *pace* Egan—content-ascription is not a necessary methodological stance. At the very best, content-ascription can play some auxiliary role in the informal presentation of a computational theory, but not within the computational model itself (Jacob 2010: 231):

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

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

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

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

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

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

Table 3

V. Concluding Remarks

In the spirit of this paper, I take contentless representationalism as an opportunity, not to reject, complement or even amend enactive anti-representationalism, but to reflect on some presuppositions and challenges of enactive anti-representationalism. Contentless representationalism invites us to retain representationalism (or at least the claim that there are subpersonal and intracranial phenomena that are naturally representational) if and only if we accept that mental representations have no content, no truth and satisfaction conditions, no reference, no intentionality, and no definitional role for cognition. The concept “mental representation” is conserved, but its meaning is deeply changed. In this sense, contentless representationalism is an alternative to anti-representationalism, which invites us to abandon the concept “mental representation” for describing or modeling natural phenomena such as neural processes, since there is no natural content at the level of intracranial and subpersonal structures. Anti-representationalism encourages the use of concepts such as indication, covariation, or correlation for defining the ac-

tivity of neural patterns. Still, for contentless representationalism, it might turn out that the concept “mental representation” refers to some kind of inner states which are currently described by anti-representationalists with these latter concepts. These states have neither content nor intentionality, but play a role in enabling the achievement of a cognitive task whose cognitive character may be defined by explanatory conventions, or by taking into account the integration of the achievement of the task in broader organism-environment interactions (if one endorses vehicle-externalism). The proponent of anti-representationalism will probably object to contentless representationalism: “WHY still go on calling these entities ‘representations’ if they do not have content, reference or intentionality?” Is not “contentless representation” a pure oxymoron, as Hutto & Myin suggest (2013: 84)? Two replies might be proposed—I see both of them as pointing towards challenging issues for enactive anti-representationalism:

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

Anti-representationalist enactivists do not seem to see that there may be scientific changes by which we *retain* concepts even though the meanings of the latter ones are changing. Or, at least, and in better words: they do not accept this possibility for the concept of “mental representation” as denoting natural, intracranial and subpersonal phenomena. But, in this case, it would be interesting to know why this possibility is refused to “mental representation” while it has been accepted for “genes” (Fox-Keller, 2002) or “atoms” (Pullman, 1998): their meanings (and the theories they figure in) have deeply changed, and yet the concepts (or minimally, the terms) have been retained³⁶. As Paul Griffiths

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

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

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

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

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

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

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

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

comes with a specific way of thinking, describing and explaining cognition, and more precisely the role of the substrate of these “mental representations”—namely, the brain—in cognitive processes. Symmetrically, if a new word should replace “mental representation”, it must come with new ways of thinking, describing and explaining cognition and brain activity. *Of course*, the enactive proponent of anti-representationalism can already provide some sketches of radically new theoretical contexts for the use of a non-representationalist vocabulary: works and developments on autopoiesis, autonomy, coupling or sense-making can help framing this context. Reconsidering what external representations are, and what cognitive jobs they can do so that we may have contentful thoughts in an anti-representationalist framework, is also an urgent task³⁸, along with a reconsideration of how the “representational hunger” of “off-line” and “abstract” cognitive tasks may be satisfied (Degenaar and Myin, forthcoming). But, as said at the very beginning of this text, enactive anti-representationalism and attempts to construct positively an enactive cognitive science are nowadays quite separated. We should only hope that this separation is very temporary and contingent. But it is also possible that this separation is the symptom of the existence of different irreducible commitments—in the enactive framework—to what is centrally required for the overcoming of classical cognitive science. Some might want to develop and to refine an explicit eliminativist stance on mental representations, while others might think that the age of representation is so over that one should not lose one’s time arguing with the past. Even if the first strategy faces some theoretical difficulties I have outlined in this paper, I believe it would be a mistake to abandon it: I do not see how the second strategy would be *a priori* immune against all forms of representationalism. And besides being parts of the same research program (namely, enactivism), these two strategies can converge on several crucial issues such as the status of meaning and content in nature (including our second nature), or the theoretical place that must be allocated to the brain in the definition and explanation of cognitive processing.

References

- Bennett, M. R. and Hacker, P. M. S. 2003. *Philosophical Foundations of Neuroscience*. Wiley-Blackwell.
- Bickhard, M. H. 2009. Interactivism: a manifesto. *New Ideas in Psychology*, 27: 85-95.
- Burian, R. M. 1985. On conceptual change: the case of the gene. D. J. Depew & B. H. Weber, *Evolution at a Crossroads – The New Biology and the New Philosophy of Science* (21-42). Cambridge (Mass.): MIT Press.

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

- Calvo Garzon, F. 2008. Towards a general theory of antirepresentationalism. *Brit. J. Phil. Sci.*, 59: 259-292.
- Chemero, A. 2000. Anti-representationalism and the dynamical stance. *Philosophy of Science*, 67: 625-647.
- Chemero, A. 2009. *Radical embodied cognitive science*. Cambridge (MA)/London: MIT Press.
- Chomsky, N. 1995. Language and nature. *Mind* 104 (413): 1-61.
- Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*, Cambridge University Press.
- Chomsky, N. 2003a. Reply to Egan. L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (268-274). Blackwell.
- Chomsky, N. 2003b. Reply to Rey. L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (274-287). Blackwell.
- Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge (MA): MIT Press.
- Clancey, W. J. 1997. *Situated Cognition: On Human Knowledge and Computer Representations*. Cambridge University Press.
- Clark, A. 1997. *Being There. Putting Brain, Body and World Together Again*, Cambridge (MA): MIT Press.
- Clark, A. 2001. *Mindware*. New York and Oxford: Oxford University Press.
- Clark, A. 2008. *Supersizing the Mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Collins, J. 2007. Meta-scientific eliminativism: a reconsideration of Chomsky's review of Skinner's *Verbal Behavior*. *Brit. J. Phil. Sci.*, 58: 625-658.
- Degenaar, J. and Myin, E. Representation-hunger reconsidered. Forthcoming in *Synthese*.
- Descombes, V. 2010. *The Mind's Provisions: A Critique of Cognitivism*. S. Schwartz transl. Princeton University Press.
- Devitt, M. 2009. On determining what there isn't. D. Murphy and M. Bishop, eds. *Stich and His Critics* (46-61). Oxford-Malden, Blackwell.
- Dreyfus, H. 2002. Intelligence without representation – Merleau-Ponty's critique of mental representation. The relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences*, 1: 367-383.
- Egan, F. 2003. Naturalistic inquiry: where does mental representation fit in? L. Antony and N. Hornstein, eds. *Chomsky and His Critics* (89-104). Blackwell.
- Egan, F. 2014 How to think about mental content. *Philosophical Studies*, 170: 115-135.
- Egan, F. 2010. Computational models: a modest role for content. *Studies in History and Philosophy of Science* (special issue on Computation and Cognitive Science), 41: 253-259.
- Ellis, R. and Newton, N. 2010. *How the Mind uses the Brain*. Open Court.

- Fodor, J. 1987. *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*. Cambridge (MA): MIT Press.
- Foglia, L. and Grush, R. 2011. The limitations of a purely enactive (non-representational) account of imagery. *Journal of Consciousness Studies*, 18 (5-6): 35-43.
- Fox Keller, E. 2002. *The Century of the Gene*. Cambridge (MA): Harvard University Press.
- Freeman, W. & Skarda, C. 1990. Representations: who needs them? J. McGaugh, N. Weinberger, and G. Lynch, eds. *Brain Organization and Memory: Cells, Systems and Circuits* (375-380). Oxford: Oxford University Press.
- Gallagher, S. 2008. Are minimal representations still representations? *International Journal of Philosophical Studies*, 16 (3): 351-69.
- Glock, H. -J. 2008. Concepts: representations or abilities? E. Di Nucci and C. McHugh, eds. *Content, Consciousness and Perception: Essays in Contemporary Philosophy of Mind* (36-61). Cambridge Scholars Publishing.
- Godfrey-Smith, P. 2004. On folk psychology and mental representation. H. Clapin, Ph. Staines and P. Slezak, eds. *Representation in Mind: New Approaches to Mental Representation* (147-162). Amsterdam: Elsevier
- Griffiths, P. and Stotz, K. 2007. Gene. D. L. Hull and Michael Ruse, eds. *The Cambridge Companion to the Philosophy of Biology* (85-102). Cambridge/New York: Cambridge University Press.
- Haselager, P., de Groot, A. and van Rappard, H. 2003. Representationalism vs. anti-representationalism: a debate for the sake of appearance. *Philosophical Psychology*, 16 (1): 5-23.
- Hutchins, E. 1995. *Cognition in the Wild*. Cambridge (MA): MIT Press.
- Hutto, D. & Myin, E. 2013. *Radicalizing Enactivism. Basic Minds Without Content*. Cambridge (MA)/London: MIT Press.
- Jackendoff, R. 1992. *Languages of the Mind*. Cambridge (MA): MIT Press.
- Jacob, P. 2010. . The scope and limits of Chomsky's naturalism. J. Bricmont and J. Franck, eds. *Chomsky Notebook* (211-234). New York: Columbia University Press.
- James, W. 1890. *The Principles of Psychology*. New York: Dover.
- Keijzer, F. 2001. *Representation and Behavior*. Cambridge (MA): MIT Press.
- Kenny, A. 1989. *The Metaphysics of Mind*. Oxford-New York: Oxford University Press.
- Lenay, Ch. and Steiner, P. 2010. Beyond the internalism/externalism debate: the constitution of the space of perception. *Consciousness and Cognition*, 19: 938-952.
- Lloyd, D. 2003. Representation, Philosophical issues about. *Encyclopedia of Cognitive Science*. NY, McMillan.
- Maturana, H. 1972. Biology of cognition. H. Maturana and F. Varela, *Autopoiesis and Cognition. The Realization of the Living* (1-58). Dordrecht – Boston: Reidel.

- Maturana, H. 1978. Biology of language: the epistemology of reality. G. Miller and E. Lenneberg, eds. *Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg (27-63)*. New York: Academic Press.
- Maturana, H. and Varela, F. 1987. *The Tree of Knowledge. The Biological Roots of Human Understanding*. Boston and London: Shambhala.
- Menary, R. 2007. *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave-Macmillan.
- Noë, A. 2004. *Action in Perception*. Cambridge (MA): MIT Press.
- O'Regan, J. K. and Noë, A. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24: 939-1031.
- Popper, K. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.
- Pullman, B. 1998. *The Atom in the History of Human Thought*. A. Reisinger (transl.). Oxford University Press.
- Ramsey, W. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press.
- Rey, G. 2003. Chomsky, intentionality, and a CRTT. L. Antony & N. Hornstein, eds. *Chomsky and His Critics (105-139)*. Oxford: Blackwell.
- Rey, G. 2003b. Representational content and a Chomskyan linguistics. A. Barber, ed. *Epistemology of Language (140-186)*. Oxford University Press.
- Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton University Press.
- Rowlands, M. 2010. *The New Science of the Mind. From Extended Mind to Embodied Phenomenology*. Cambridge (MA)/London: MIT Press.
- Sellars, W. 1956. Empiricism and the philosophy of mind. H. Feigl and M. Scriven, eds. *Minnesota Studies in the Philosophy of Science, Volume I: The Foundations of Science and the Concepts of Psychology and Psychoanalysis (253-329)*. University of Minnesota Press.
- Shapiro, 2014. Review of *Radicalizing Enactivism. Basic Minds without Content*, Mind, forthcoming.
- Sprevak, M. 2013. Fictionalism about neural representations. *The Monist*, 96 (4): 539-560.
- Steiner, P. 2010. The bounds of representation. A non-representationalist use of the resources of the model of extended cognition. *Pragmatics & Cognition*, 18 (2): 235-272.
- Steiner, P. 2011. Enaction, pragmatisme et minimalisme représentationnel. N. Bault, V. Chambon, N. Maïonchi-Pino, F.-X. Pénicaud, B. Putois and J.-M. Roy, eds. *Peut-on se passer de représentations en sciences cognitives? (191-210)*. Bruxelles-Paris: De Boeck
- Steiner, P. 2013. Les radicaux de la cognition. A propos de *Radicalizing Enactivism. Basic Minds without Content (2013)*. *Intellectica*, 59, 2013/1: 299-308.
- Steiner, P. 2014. A problem for representationalist versions of extended cognition. *Philosophical Psychology*, forthcoming.

- Steiner, P. 2014b. The delocalized mind. Judgements, vehicles, and persons. *Phenomenology and the Cognitive Sciences*, 13 (3): 437-460
- Stewart, J. 1996. Cognition = Life: implications for higher-level cognition. *Behavioural Processes*, 35: 311-326.
- Stewart, J., Gapenne, O. and Di Paolo, E. A., eds. 2010. *Enaction. Towards a New Paradigm for Cognitive Science*. Cambridge (MA): MIT Press.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science: The Case against Belief*. Cambridge (MA): MIT Press
- Stich, S. 1996. *Deconstructing the Mind*. Oxford-New York: Oxford University Press.
- Sutton, J. 2010. Exograms and interdisciplinarity: history, the extended mind, and the civilizing process. R. Menary, ed. *The extended mind* (189-225). Cambridge (MA)/London: MIT Press.
- Thelen, E. and Smith, L. B. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge (MA): MIT Press.
- Thompson, E. 2011. Replies to commentaries. *Journal of Consciousness Studies* 18 (5/6): 76-223.
- Van Gelder, T. 1995. What might cognition be if not computation ? *Journal of Philosophy*, 92: 345-381.
- Varela, F. 1989. *Autonomie et connaissance. Essai sur le vivant*. Paris: Seuil.
- Varela, F., Thompson, E. & Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge (MA): MIT Press.
- Webb, B. 1994. Robotic experiments in cricket phonotaxis. D. Cliff, P. Husbands, J.-A. Meyer and S. W. Wilson, eds. *From animal to animats 3. Proceedings of the third international conference on simulation of adaptive behaviour*. Cambridge (MA): MIT Press.
- Wheeler, M. 2005. *Reconstructing the Cognitive World. The Next Step*. Cambridge (MA): MIT Press.
- Wilson, R. 2004. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge: Cambridge University Press.
- Wilson, R. and Foglia, L. 2011. Embodied Cognition. *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition). Edward N. Zalta, ed. URL = <<http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/>>.
- Zahidi, K. 2014. Non-representationalist cognitive science and realism. *Phenomenology and the Cognitive Sciences*, 13 (3) : 461-475.