Is radically enactive imagination really contentless?

Marco Facchin [**corresponding author**]
Istituto Universitario di Studi Superiori IUSS Pavia
Department of Human and Life Sciences; Pavia, Italy
Palazzo del Broletto, Piazza della Vittoria n. 15, 27100, Pavia
marco.facchin@iusspavia.it
https://orcid.org/0000-0001-5753-9873

Ruhr Universität Bochum
Fakultät für Philosophie und Erziehungswissenschaft
Institut für Philosophie II; Bochum, Germany
Universitätsstr. 150, GA3/29, 44780 Bochum

Abstract:

Radical enactivists claim that cognition is split in two distinct kinds, which can be differentiated by how they relate to mental content. In their view, basic cognitive activities involve no mental content whatsoever, whereas linguistically scaffolded, non-basic, cognitive activities constitutively involve the manipulation of mental contents.

Here, I evaluate how this dichotomy applies to imagination, arguing that the sensory images involved in basic acts of imaginations qualify as vehicles of content, contrary to what radical enactivists claim. To argue so, I leverage what has appropriately been dubbed a "compare to prototype" argument. Hence, I will first identify, within the enactivist literature, the general functional profile of a vehicle of content complying with the austere standard of contentfulness radical enactivists adhere to. Provided such a profile, I will show, relying on a mixture of reasoning and empirical evidence, that basic sensory images satisfy it, and thus that they can rightfully be identified as vehicles of content. This, I claim, provides a sufficient reason to identify the sensory images involved in basic acts of imagination as vehicles of content, thereby denying that basic imagination does not involve mental content.

Keywords: Radical Enactivism; Enactive Imagination; Mental Content; Predictive Processing

READCUBE: https://rdcu.be/cdtqT

Formatted in comic sans to be more easily readeable by people suffering from dyslexia.

Declarations:
Founding: This work has been funded by the PRIN Project "The Mark of Mental" (MOM), 2017P9E9N, active from 9.12.2019 to 28.12.2022, financed by the Italian Ministry of University and Research.
Conflict of Interests/competing interests: The author declares no conflict of interests.
Availability of data and Material: Not applicable.
Code availability: Not applicable.
Authors' contribution: Marco Facchin is the sole author of the paper.
Acknowledgments: This early version of this works have been presented in a number of conferences, most notably the SILFS Postgraduate conference at Urbino University and the Pretend Play and E-Cognition conference at the University of Antwerp. I would like to thank the audience of both of these talks for the helpful comments.

1. Introduction

Enactivist philosophers typically understand cognitive processes as sensorimotor interactions enacting mental content. They claim that, by interacting sensomotorically with a target, an agent comes to establish a meaningful perspective onto it: whether it should be loved or feared, avoided or proactively sought (e.g. Varela, Thompson and Rosch 1991; Thompson 2007).

Not every enactivist agrees with this picture. Specifically, radical enactivists argue that most of an agent's cognition involves no content at all (Hutto and Myin 2013; 2017).[1] In their view, cognition is split in two distinct kinds. Basic forms of cognition involve no content and, accordingly, should be explained without referencing mental content. Conversely, non-basic

[1] The opposition between radical and "non radical" enactivism is partially artificial: "non radical" enactivism is a multi-faceted movement (see Ward, Silverman and Villalobos 2017). Yet, since radical enactivism is singled out from other forms of enactivism by the claim that there is contentless cognition, their sharp opposition seems to me warranted.

forms of cognition are constitutively content-involving and revolve around the manipulation of public symbols, especially linguistic ones.

Here, I attack the radical enactivists' conception on cognition, focusing on basic (i.e. contentless) imagination.[2] I aim at providing a sufficient reason to deny that basic acts of imaginations are contentless. To do so, I leverage what Gładziejewski (2016) has appropriately dubbed a "compare to prototype" argument; showing that basic sensory images fit the radical enactivists' prototype of a vehicle of content. This, I claim, provides a sufficient reason to identify basic acts of imagination as vehicles of content, thereby denying that they are contentless.[3]

The essay is structured as follows. The next section briefly introduces the framework of radical enactivism, focusing on basic imagination. Section three introduces "compare to prototype" arguments, and provides a functional profile of vehicles of content complying with the radical enactivists' standards on contentfulness. Section four shows that basic acts of imagination satisfy that profile, and hence that they qualify as vehicles of content. A brief concluding paragraph closes the paper.

2. The framework of radical enactivism: basic and non-basic imagination

As hinted in the introduction, what differentiates radical enactivism from its "non radical" counterpart is the claim that some of an agent's cognitive interactions with the environment do not involve mental content. More precisely, radical enactivists argue that there are two distinct kinds of cognitive interactions, to be distinguished by how they relate to mental content (Hutto and Myin 2017: 134).

---

[2] Throughout the essay, I will use "basic" and "contentless" interchangeably. The same holds for "non-basic", "contentful" and "content-involving".

[3]Here, following Hutto and Myin (2017) I will use "imagination" and "sensory images" interchangeably. However, as recently argued in Arcangeli (2019) and Wiltsher (2019: 13-17), this usage is problematic. In fact, sensory images can also be involved in non-imaginative acts. For instance, I can conjure a sensory image to *recall* how my mother used to look in her twenties. Many thanks to an anonymous reviewer for having noticed this issue.

This distinction can be made more precise by examining content-involving cognition first. Radical enactivists hold that only enculturated and socially scaffolded human beings are capable of non-basic cognition, and thus of contentful thought (Hutto and Satne 2015; 2017; Myin 2020). This is because, on their account, non-basic cognitive activities constitutively involve the mastery of sociocultural practices, which regiment the manipulation of content-involving public symbols, especially linguistic ones (Hutto and Myin 2017: 121-146). On this view, it is only within the background provided by such practices that what an agent says and does can be semantically evaluated by how it meets its conditions of satisfaction (e.g. truth or accuracy conditions), with which radical enactivists identify mental content (see Hutto and Myin 2013: ix-xx; Hutto 2015: 73). Radical enactivists hold that mental content is always located *outside* our brains, within the culturally regimented acts of an agent (Myin 2020). As a result, each cognitive phenomenon that does not take place against such a culturally established backdrop is a case of contentless cognition. Contrary to the popular belief, radical enactivists argue that almost all cognition is contentless (Hutto and Myin 2017: 135).[4]

Notice that radical enactivists do not claim that only "higher" forms of cognition (e.g. mathematical cognition) involve content. In fact, radical enactivists argue that all forms of cognition are contentful, if appropriately regimented by an agent's mastery of cultural practices (Myin 2020). For instance, basic forms of memory can be turned into non-basic ones by mastering narrative practices (Hutto and Myin 2017: 206-214). Similarly, basic forms of perception are transformed in non-basic ones by engaging in perceptual judgments (*ibidem*: 171). Thus, albeit distinct in kind, basic and non-basic forms of cognition do not indicate different cognitive processes. The *same* cognitive process can be either basic or non-basic, depending on its embedding in a cultural context. Or so radical enactivists argue.

[4] I'm deeply thankful to the anonymous reviewer who noticed that this point needed further clarification.

Applied to imagination, the distinction between basic and non-basic cognition articulates as follows (Hutto 2015; Hutto and Myin 2017: 177-201). Non-basic imaginative acts are hybrid acts, in which a basic, *per se* contentless, sensory image is conjoined with an imaginative attitude (Langland-Hassan 2015).[5] These attitudes provide basic sensory images with an appropriate "linguistic component" (Hutto and Myin 2017: 193), which determines how the image is used, and thus whether the relevant conditions of satisfaction provided by the imaginative attitude obtain. Radical enactivists clarify what they mean proposing the following example (*ibidem*: 189-190). Suppose a person conjures the sensory image of a big silver arch, and uses it to judge how the Arc de Triomphe looks like. In this case, the person's imagination fails her (the Arc de Triomphe is not a big silver arch). But this failure is entirely due to the linguistic context in which the sensory image was embedded: namely, that of judging how the Arc de Triomphe looks like. In fact, the very same sensory image would constitute an imaginative success if embedded in a different linguistic context (if, say, the person is trying to judge how the Arc de Triomphe would look like, were it made out of silver). As I understand it, this example aims to show that the conditions of satisfaction of a non-basic imaginative act do not belong to the sensory image *per se*, but rather to the linguistic component the sensory image is conjoined with.

Basic acts of imagination consist in the production of the same sensory images, *minus* the linguistic embedding. In a non-linguistic context, these sensory images directly control an agent's online sensorimotor interactions with the environment, as they are needed to "inform anticipatory behavior and to guide, or at least adjust, any intelligent engagements" (Hutto and Myin 2017: 196). To make their view more vivid, radical enactivists point at the case of hominid tool-making (*ibidem*: 193-194), speculating that these basic sensory images could

[5] To be clear: Langland-Hassan (2015) takes sensory images to be contentful. More precisely, he argues that these images have contents akin to indefinite descriptions: descriptions beginning with "a" or "some" (see Langland-Hassan 2015: 13). These contents are then further specified by the imaginative attitude to which the image is conjoined. Conversely, on Hutto and Myin's view, the sensory image has no content. The relevant content is entirely carried by the imaginative attitude. Thanks to an anonymous reviewer for having encouraged me to clarify this point.

have enabled hominids to exert their tool-making skills by guiding their online interactions with their tools even in an a-linguistic (or protolinguistic) context.

But what are these basic sensory images? If I understand radical enactivists correctly, the answer to this question is provided by their account of perception (and action), which is a contentless interpretation of predictive processing (Hutto and Myin 2017: 57-74; 147-176). According to this framework[6], perception is a top-down process, by which the perceiver tries to anticipate the incoming sensory inputs, leveraging an acquired body of knowledge (called a generative model) about how sensory inputs are generated by the environment (see Clark 2014; 2016). These anticipations take the form of an endogenously generated flow of "virtual" sensory inputs, traveling through the brain from associative to primary sensorimotor cortices (see Mesulam 2008). Crucially, this endogenously generated flow of virtual inputs is held responsible for the production of "transient percepts", such as the ones experienced when perceiving an omission (see Adams *et al.* 2013) or in more mundane cases of visual imagery (Clark 2012; 2014: 18; 2016: 94-98).[7]

More precisely, the cognitive apparatus predictive processing describes works as follows. In perception, the agent generates "from the top-down" an endogenous flow of virtual inputs, which are then contrasted with the ones actually delivered by the environment. The mismatch between the two generates a "bottom-up" internal signal (known as prediction error), which

---

[6] Notice: here the focus is shifting from radical enactivism to predictive processing. As predictive processing is (strictly speaking) a neurocomputational theory, I will introduce it using a computational lexicon. For instance, I will say that brains try to anticipate the incoming sensory *inputs* using stored *knowledge*. Radical enactivists, however, would prefer to talk about embodied anticipations generated by a self-organizing system in virtue of the system's history of interactions (see Hutto and Myin 2017: 57-67; Hutto 2018). In a similar spirit, they would substitute "predictions" (and the like) with "robust correspondences between neural and worldly dynamics". Importantly, predictive processing can also be articulated in a more dynamicist vocabulary (see Tani 2016). Many thanks to the anonymous reviewer who noticed that this point needed to be clarified further.

[7] It should be noted that radical enactivists focus *only* on motor emulators (i.e. forward models) when discussing basic imagination (Hutto and Myin 2017: 197-199). Given the rough computational similarity of motor emulators and generative models (see Pickering and Clark 2014), I interpret radical enactivists as endorsing a predictive processing account of basic imagination. This seems the most charitable interpretation of their position, as it bolsters the internal consistency of radical enactivism. It also boosts its explanatory power: albeit motor emulators can account for motor imagery, it is far from clear how motor emulators *alone* could account for visual imagery. Thus my argument presupposes that radical enactivists endorse a predictive processing account of basic imagination, albeit radical enactivists are not explicit on this point. Importantly, however, if they do not endorse a predictive processing account of basic imagination, then their account seems to lack in explanatory power, as it seems unable to account for visual imagery. And this would be a serious flaw in its own right.

the perceiver leverages to form new, more apt, virtual inputs, which will minimize the incoming error signal (Clark 2016: 13-52). Perceiving, thus, appears as a process in which one aligns one's own endogenously generated virtual inputs to the ones delivered by the environment. Importantly, the very same mechanism is also responsible for action. According to predictive processing, embodied actions literally consist in the enactment of one's virtual, endogenously generated, sensory expectations (see Adams, Shipps and Friston 2013).[8] In sum, these virtual inputs seem responsible for the appearance of imaginary percepts in different modalities, and are ideally suited to "inform anticipatory behavior and to guide, or at least adjust, any intelligent engagements" (Hutto and Myin 2017: 196). As such, they seem to be what radical enactivists call basic sensory images.

This is admittedly an highly schematic and incomplete presentation of predictive processing. Yet it should be sufficient to clarify what are the basic images invoked by the radical enactivists' conception of imagination. Basic sensory images are nothing but the endogenously generated flow of virtual inputs, that, according to predictive processing, services online perception and action. As such, basic imagination is interwoven with online sensorimotor engagements, just as required by the extensive, always interaction-involving, conception of the mind radical enactivists defend (see Hutto and Myin 2013: 135-153; Hutto, Kirchhoff and Myin 2014).

Having succinctly presented the radical enactivists' conception of basic imagination, I move to introduce "compare to prototype" arguments.

3. "compare to prototype" arguments

Ramsey's (2007) book *Representation Reconsidered* significantly shifted the way philosophers of cognitive science approach the problems of representation and content. Ramsey's analysis begins by noticing that, according to the current practice of cognitive

[8] In other words, perception and action are generated by the same underlying process of error minimization. They differ only in direction of fit.

science, virtually every structure mediating between sensory input and motor output would qualify as a representation. Ramsey finds this liberal understanding of representations problematic, as it *trivializes* the concept. He argues that, if everything counts as a representation in some sense, then identifying cognitive structures as representations brings about no explanatory advantage, as representations are no longer a distinct kind of posits doing some unique explanatory work within cognitive theories. On Ramsey's view, only certain specific kinds of posits should be categorized as representations: namely, the posits that play the functional role of stand-ins for their targets.

Here is where "compare to prototype" arguments[9] come into play, to identify these posits. These arguments start by identifying a representational prototype: a class of entities we pre-theoretically and uncontroversially identify as representations. The prototype is then examined, determining its functional profile: the way in which it "stands in" for what it represents. Lastly, some allegedly representational structure of interest is examined, to see whether it satisfies the functional profile abstracted from the prototype. If it does, then the allegedly representational posit under examination really is a representation, whose specific representational role is now understood: it functions as a representation by functioning as the given representational prototype (e.g. a map, a sentence, a symbol etc). Arguments of this kind have been widely leveraged to determine the status of representational posits, yielding both representationalist and anti-representationalist verdicts (e.g. Orlandi 2014; Gładziejewski 2016; Dołega 2017; Downey 2018); and they are in general considered to be sufficient to determine the representational status of a posit.

Here, I leverage one such argument, to claim that basic sensory images, as radical enactivists depict them, fit the radical enactivists' prototype of a vehicle of content. Hence, contrary to radical enactivists' claim, they should be conceived as contentful. Clearly, to do so, I need the general prototype of a vehicle of content.[10] Moreover, I need to abstract the

[9] The naming is due to Gładziejewski (2016).

[10] Here, I use "vehicle of content" in an uncommitted way, to designate whichever concrete particular bears a

relevant functional profile from it. Luckily for me, however, the relevant functional profile is provided by the enactivist literature, in the form of Rowlands' (2006; 2018) five functional features that identify contentful vehicles. But before examining them, let me clarify a point.

In standard "compare to prototype" arguments, one starts with the prototype, then abstracts its functional profile, and, lastly, checks whether the posit under examination satisfies the functional profile thus abstracted. Here, the ordering will be different: I will start with the functional profile, and then check whether it captures the relevant features of the radical enactivists' prototype of a vehicle of content; namely, a linguistic utterance. This will be done in this section. In the next, I will check whether basic sensory images satisfy the relevant profile. Notice that albeit the ordering is non-standard, the difference between this and standard "compare to prototype" arguments is minimal. In both cases what matters is whether the posit under scrutiny satisfies the relevant functional profile. The prototype itself plays a modest role: it is used only to show, in an intuitively perspicuous manner, that the relevant functional profile is adequate to capture genuine representations.[11] Here, the prototype will play the same role: it will intuitively show that the relevant functional profile captures what, *according to radical enactivism*, are genuine vehicles of content. Moreover, as in standard "compare to prototype" arguments, here too the whole force of the argument hinges on whether the posit under scrutiny satisfies the relevant functional profile. Hence, it seems to me that the most relevant change in my non-standard "compare to prototype" argument is just the order in which the ingredients will be put on the table. The execution of the dish (so to speak) will be identical.

Now, the functional profile (Rowlands 2006; 2018). An item R is a vehicle of content about a target state of affairs S only if:

---

metaphysically privileged relation to mental contents. This usage is consistent with the enactivist literature (e.g. Hurley 1998; Menary 2007).

[11] It would be a mistake, for instance, to check whether a given posit is intuitively similar to the prototype: the so-called "cognitive map" in the rat hippocampus surely does not resemble a cartographic map. Yet, "cognitive maps" are the paradigmatic case of a representational posit recognized as such by "compare-to-prototype" arguments (e.g. Ramsey 2016).

(1) R carries information about S; &

(2) R has either the proper function of tracking S or the proper function of enabling the tracking of S; &

(3) R can misrepresent S; &

(4) R belongs to a wider representational framework; &

(5) R is decouplable from S

It is impossible, due to space limitations, to individually justify (1) to (5) here.[12] Yet these five conditions should be familiar enough, as they have been a staple in the philosophical discussion on content since at least the 80's. For the same reason, they need little unpacking.

Starting with (1), the relevant notion of information at play is that of covariation (Hutto and Myin 2013: 63-71; 2018: 99-102). Succinctly, R carries information about S only if the states of R and S reliably covary over some range of states, in a way such that observing R being in a state *r*' makes one more confident of S being in a state *s*'.

Covariance relations satisfying (1) are ubiquitous in nature (e.g. the level of the sea covaries with the position of the moon). But (2) imposes a significant restriction on (1). In fact, (2) does not just require the covariance relation in (1) to be non-accidental. It also imposes that R has to (normatively speaking) covary with S. Rowlands (2006; 2018) unpacks the normative component of (2) appealing to proper functions (Millikan 1984). Radical enactivists find proper functions to be somewhat restrictive (Hutto and Myin 2017: 117). If I understand them correctly, they suggest to identify as a function of an item whatever output the item produces robustly, in virtue of some selection process selecting items of that class on the basis of the output they produced. These selection processes might be various in nature: natural selection surely is one, but individual and social learning count too (Hutto and Myin 2017: 117). To my knowledge, radical enactivists do not substantially expand their notion of function, and so this rough characterization must do for now.[13]

---

[12] See Rowlands (2006) for a book-length treatment.

[13] It is of course possible to appeal to more defined, but still liberal, notions of function (e.g. task functions, see Shea

Luckily, (3) is far more defined. It can be unpacked in reference to (2) and (1). A vehicle of content is supposed to accurately depict its target. Thus, a vehicle of content misrepresents everytime it fails to carry out its proper function. Given that the relevant proper function (2) imposes on vehicles of content is that of tracking (i.e. covarying with) their targets, it follows that vehicles of content misrepresent everytime (1) fails to obtain.

This might sound paradoxical. The functional profile is satisfied if, and only if, (1) to (5) are satisfied in conjunction. Yet, (3) requires that (1) might not be satisfied. The paradox, however, is illusory. For (1) and (2) refer to vehicles as types, whereas (3) to (5) refer to vehicles as tokens.[14] A given type of vehicle satisfies (1) when vehicles of that type reliably and predictably covary with their targets, thus carrying information about the state of their targets. Similarly, the relevant notion of function applies to types of vehicles, as selection processes operate over classes of items. A single token vehicle, however, might fail to carry out its proper function, thus failing to satisfy (1). Compare: the position of pc cursors (notice the plural) is supposed to covary with the position of the mouse. However, if a computer malfunctions, a single pc cursor might fail to do so. Similarly, different types of vehicles are supposed to keep track of different targets, albeit sometimes individual (token) vehicles might fail to do so.

The pc cursor example is also functional to address one possible complaint voiced by radical enactivists. The worry is basically the following: as the pc cursor example shows, failures of tracking are not *eo ipso* failures in representing (hopefully, no one holds that pc cursors represent mouses). Indeed, radical enactivists have discussed this point at length, providing a contentless account of tracking and a contentless account of functions (Hutto and Myin 2017: 104-114; see also Hutto 2013). Thus, failures in tracking should not be equated with failure in representing, not even when the failing tracker has the function of tracking its target. Agreed. Yet, albeit I agree with radical enactivists that satisfying (1) to (3) is not

2018). But radical enactivists do not endorse them, and I'm playing by their rules.

[14] Many thanks to an anonymous reviewer that noticed my original formulation led to a paradox.

sufficient for misrepresentation to occur, I am not claiming that satisfying (1) to (3) is sufficient for misrepresentation to occur: (4) and (5) must be satisfied too.[15] Notice, in fact, that if a candidate vehicle were not able to satisfy also (4) and (5), the candidate vehicle would not qualify as a vehicle of content according to the proposed functional profile. Hence, it would not be deemed able to represent and misrepresent. It would, however, be deemed able to track and mistrack. Given that here (1) to (5) are required in conjunction, spelling out (3) in terms of misrepresentation strikes me as intuitively correct. A radical enactivist might wish to re-read (3) substituting every occurrence of "misrepresent" (and the like) with "mistrack" (and the like). But these substitutions would be substitutions *salva veritate* and, as such, they would not challenge my argument.

Point (4) highlights the fact that vehicles of content typically belong to a wider class of other such vehicles (e.g. Haugeland 1991), which can productively be combined to yield new, more complex, vehicles. Rowlands (2006: 177-182) however makes clear this constraint is pretty liberal. In fact, (4) does not require (nor excludes) that vehicles of content must exhibit a compositional structure similar to that of natural languages. All (4) requires is either that vehicles of content belong to a finite set of vehicles that can be freely recombined so as to yield more complex vehicles of content, or that they are the upshot of such a recombination.

Lastly, condition (5) requires R to be decouplable from S. The standard notion of decouplability is that of an absence of causal contact (e.g. Rowlands 2006: 163-164; Chemero 2009: 47-66; Gładziejewski 2015). In other words, R is decoupled from S when there is no mechanical energy transfer from S to R (or *vice versa*) such that S can systematically determine the state of R (or *vice versa*). The point should be at least descriptively familiar: a

[15] For the same reason, I'm not suggesting that robust tracking (i.e. satisfying (1) to (3)) alone is sufficient for something to depict accurately (recall the pc cursor example). Many would in fact notice that there are a number of non-representational structures able to track some target robustly (Hutto and Myin 2017; see also Ramsey 2007: 118-250).

picture need not constantly be in causal contact with the depicted item, nor a sentence needs to be.

Notice two important features of conditions (1) to (5). Firstly, condition (1) to (5) are all functional constraints on vehicle properties: they all point to specific features of concrete particulars encoding content. It is the vehicle that (1) covaries with a target, (2) has the function to do so, (3) can fail to do so, (4) belongs to a finite set of other particulars or is the upshot of the combination of particulars belonging to such a set and (5) can be in no causal contact with the target with which it covaries. Secondly, conditions (1) to (5) *need not* be (albeit they might be) what determines the content of a vehicle. I will expand this point in the conclusion of the next section. For now, it is sufficient to notice that, within the framework Ramsey proposes, questions about content determination and questions about representational function are squarely separated (see Ramsey 2007; 2016). I will capitalize upon this feature of the framework.

Are (1) to (5) jointly sufficient to identify vehicles of content? I believe they are. Consider a paradigmatic vehicle of content, such as the banal utterance "Mary is left of John". On the supposition that the speaker is honest, it jointly satisfies (1) to (5). It carries information about Mary, John and their relative position, as (1) requires: an honest speaker would not have uttered "Mary is left of John" weren't Mary left of John. The utterance and the state of affairs it represents, thus, appear to covary across at least close counterfactual scenarios.[16] The utterance also seems to have some specific functions. It surely seems to have the function to indicate that Mary is left of John, and may have other functions. It might, for instance, have the function of generating a specific belief in the hearer, allowing her to track a distal state of affairs. In fact, speaker, utterance and hearer have been reciprocally fit, through the processes of enculturation and learning, in a way such that

[16] Notice that, on Dretske's account, this kind of counterfactual dependence is what *constitutes* the information carrying status of an item (Dretske 1988: 56-57). Since radical enactivists typically point at Dretske when discussing information, I suppose that they agree.

simple indicative statements are supposed to prompt certain beliefs in their hearers. Hence (2) seems to obtain. Utterances can surely misrepresent, and they boast the combinatorial structure of human natural languages, nicely vindicating (3) and (4). Lastly, utterances can be decoupled from the states of affairs they target. That statement might, for instance, be uttered by looking at a photograph of Mary and John, or simply be uttered by a speaker for no apparent reason (other than making a philosophical point).

Importantly, a banal utterance such as "Mary is left of John" qualifies as a vehicle of content even according to the austere standard of contentfulness radical enactivists adhere to. Utterances are in fact public linguistic symbols, the manipulation of which is carried out by an enculturated agent according to the cultural and social norms of a community of speakers. Hence conditions (1) to (5) identify vehicles of content that radical enactivists would recognize *as such*. Thus, conditions (1) to (5) provide a functional profile that meets the radical enactivists' standards.

This concludes the presentation of the functional profile. Notice that it sets a pretty low bar against which to test radical enactivism. The profile itself comes from the enactivist literature, and captures what radical enactivists themselves take to be the prototype of a vehicle of content, namely a linguistic utterance. Moreover, as detailed above, "compare to prototype" arguments yield both representationalist and anti-representationalist verdicts, so they are not inherently biased in favor of any outcome. Thus, if really basic sensory images are contentless, they should violate (1) to (5) blatantly.

In the next section, I will argue that basic sensory images, as radical enactivists conceive them, satisfy conditions (1) to (5) in conjunction, providing a sufficient reason to identify them as vehicles of content, contrary to what radical enactivists argue.

4. Applying the profile to basic sensory images

To start, recall how radical enactivists conceive basic sensory images. Radical enactivists conceive basic sensory images as the top down, endogenously generated, flow of virtual sensory inputs that predictive processing posits to account for online perception and action. My task is now that of showing that these virtual sensory inputs satisfy (1) to (5), and thus that they fit the functional profile of a vehicle of content provided above. Showing this is sufficient, if "compare to prototype" arguments are on the right track, to demonstrate that basic images are vehicles of content, and thus that basic imagination involves content, contrary to what radical enactivists claim.

Consider condition (1). There are strong intratheorethical reasons to hold that basic sensory images carry information about their targets. Indeed, if the predictive processing account of perception and action is correct, they must (see Hohwy 2013: 41-58 for a nice illustration). I will not, however, pursue this line of argument here, as radical enactivists will likely find it unconvincing (see Kirchhoff and Robertson 2018). Yet, aside from information-theoretic considerations, there are other reasons to hold that basic images covary with their targets as (1) requires. Recall the succinct account of predictive processing sketched above: to correctly perceive a target, the perceiver must align its endogenously generated flow of virtual sensory signals with the sensory signals actually generated by the target, minimizing prediction error.

But for prediction error to be minimized, virtual and actual sensory inputs must match. This already introduces covariance across counterfactual scenarios: to correctly perceive an object, the perceiver would have to generate a virtual flow of input *v** instead of *v*, had the perceived object delivered an input *i** instead of *i* (Clark 2016: 92).[17] Covariance, however, is by no means limited to counterfactual scenarios. Indeed, it holds across time. For instance, if predictive processing is correct, to visually track an object just *is* to endogenously generate the virtual inputs the object will actually deliver in its next position (see Adams, Perrinet and

[17] The same holds for imagination: to correctly imagine an object, a subject would have to generate a flow of virtual input *v** instead of *v*, where the imagined object is *i** instead of *i*.

Friston 2012). Of course, on this account, the covariance relation is not established exclusively by the effect the perceived target has upon the perceiver's sensorium: it is also due to the perceiver's endogenous generation of virtual sensory signals. This is a significant difference between predictive processing and more canonical, bottom-up, accounts of perception. Yet, the fact that the covariance relation is partially endogenously generated within the perceiver does not make it less of a covariance relation. Therefore, (1) obtains.

Consider now condition (2). Surely the covariance relation between the endogenously generated virtual sensory input and the one actually delivered by the environment is nonaccidental. But have these virtual sensory signals the function of tracking the environment? It seems they have, especially under the more relaxed notion of function radical enactivists appear to endorse.

To see why this is the case, consider that prediction error offers a constant feedback signal that the perceiver uses to improve her generation of virtual sensory input overtime, with the aim of minimizing prediction error (Hohwy 2013: 48-51; Clark 2016: 17-19). Streams of endogenously generated virtual sensory signals are constantly selected based on how they robustly manage to produce an outcome; namely prediction error minimization. But, as clarified above, error minimization entails covariance. Therefore, these streams of virtual inputs are selected for how they covary with their targets. On longer timescales, this process corresponds to individual learning (see Friston *et al.* 2016), and radical enactivists state explicitly that individual learning is sufficient to determine functions (see Hutto and Myin 2017: 117). But the very same selection process also operates on shorter timescales. In fact, different and competing flows of virtual sensory input are constantly generated in parallel (e.g. Friston 2005: 823). But only one of them can determine the agent's perceptual contact with the world and the subsequent course of action (Hohwy 2013: 207-223; Clark 2018). Hence, even in shorter timescales, different streams of virtual sensory inputs are contrasted, to select the one that best minimizes prediction error. Thus, even in (relatively)

short timescales, virtual sensory signals are selected[18] for their ability to produce an output (i.e. error minimization), and can thus be assigned a function according to the liberalized notion of function radical enactivists seem to endorse.

Are basic sensory images able to misrepresent[19] as (3) requires? I believe this can be shown indirectly. Recall, to start, the analysis of (3) provided above. According to it, misrepresentation occurs whenever a token vehicle fails to keep track of its target, violating the covariance relation in (1). This entails that a vehicle of content that misrepresents carries no information about its intended target, as it violates (1). Being able to violate (1) just is being able to satisfy (3).

Consider now that the kinematic[20] of actions seems to carry information about their targets (e.g. Ambrosini *et al.* 2013). This has been established by comparing the gaze allocation of experimental subjects looking at two kinds of videos, both displaying an actor reaching for one of two objects of different sizes (Ambrosini, Costantini and Sinigaglia 2012). In the first kind of videos, the actor was shown reaching one of the two objects to grasp it. Since the two objects were of different sizes, the actor was forced to preshape its hand differently depending on which object was the target of the action. In other words, to fluently grasp the target, the actor had to position its fingers differently before making actual contact with the object, and in a way that depended on the size of the object to be grasped. This allowed the subjects to proactively identify the target of the action, and to gaze upon it *before* the actor's hand grasped it. In the second kind of video, the actor merely touched the object, and so no preshaping of the hand was required. In this case, participants were not able to proactively identify the target of the action.[21] These data suggest that the relation of covariation between the position of the actor's finger and the size of the object

[18] Notice the usage of a "selectionist" vocabulary is fully justified in predictive processing. In fact, at least sometimes, an agent has to *select* a model (among competing ones) of how sensory signals are generated (see FitzGerald, Dolan and Friston 2014: 2).

[19] Again, here the radical enactivist might wish to read "mistrack" instead of "misrepresent". That would not affect the point I'm trying to make, as clarified above.

[20] Here, I use "kinematic" as it is used in physics, to denote the movement of bodies (in this case, hands) through space.

[21] In fact, they resorted to a default strategy, looking at the most visually salient target first.

carried information about the target of the action, allowing the subjects to proactively identify it.

Notice that "carrying information" is a transitive relation: if A carries information about B, and B does so about C, then A carries information about C. Now, if the predictive processing story about action is on the right track, bodily movements merely enact flows of endogenously generated basic sensory (mainly proprioceptive) images. Therefore, if the kinematic of action carries information about its target, and if that kinematic just is the enaction of endogenously generated basic images, then, by transitivity, these images carry information about the target of the action.

Consider now what happens when a subject sees a *failed* reach-to-grasp action, as in (Costantini, Ambrosini and Sinigaglia 2012). Just as in the experiment above, subjects were required to look at two kinds of videos, displaying an actor reaching for one of two targets, either to touch or to grasp it. Crucially, however, in some videos the objects were out of the actor's reach, and the actor's action failed. In these cases, the informational effect of hand preshaping failed to manifest. When the target was out of reach, subjects not only took the same time to identify the target in reach-to-touch and reach-to-grasp conditions; they were also slower than they were in the reach-to-touch condition for targets in reach. The information carried by the kinematic of the actions simply disappeared. But, just as before, that kinematic is just the enaction of a flow of endogenously generated virtual sensory input. So, if the kinematic of an action can fail to carry information about its target, then the virtual sensory inputs can too. Hence, they can violate (1). But being able to violate (1) amounts to satisfying (3).

These virtual sensory inputs seem also to have a combinatorial nature, as required by (4). Consider, for instance, the *f*MRI study detailed in (Buccino *et al* 2004a). Musically naive subjects were asked to lay in the scanner, and then look at a video of a hand playing a chord on a guitar neck. After having seen the video, subjects had to prepare themselves to re-enact

the same chord. Subjects were then required to actually enact it (on a fake, scanner-compatible, guitar neck), while still in the scanner, as a go signal was administered. This procedure identified two clusters of activation that are relevant for the discussion at hand.

The first is an activation of the fronto-parietal mirror circuit, detected when the participants were observing the chord being played. This activation of the mirror neuron circuit is not *in itself* surprising (participants were in fact observing an action), but needs careful handling, as one can "mirror" only movement belonging to one's own motor repertoire (see Buccino *et al* 2004b; Rizzolatti and Sinigaglia 2006). As participants were musically naive, they were unable to mirror the whole "playing a chord" action. Hence, researchers concluded they must have mirrored single finger movements. In other words, they concluded participants decomposed the whole action in its *constituent parts*, and mirrored the parts thus identified.

The second relevant (for the limited purposes of this essay) pattern of activation was identified in the preparation condition (i.e. the time between the end of the video and the administration of the go signal), and consisted in an activation of the same fronto-parietal mirror circuit plus an activation of the left dorsolateral prefrontal cortex (DLPFC). As the DLPFC is typically associated with executive functions and working memory, the experimenters interpreted its activation as subserving the orderly *recombination* of the single elementary movements previously identified.[22]

Now, according to the predictive coding account of mirror neurons (see Kilner, Friston and Frith 2007; Friston, Mattout and Kilner 2011), what mirror neurons recognize (and generate) are kinematic expectations about the course of action. And, as argued when discussing the previous point, these expectations are just the downward, self-generated, endogenous flow of virtual sensory inputs. Hence, on the predictive processing account radical enactivists

[22] And in fact, the involvement of the DLPFC decreases when the subjects *can* mirror the whole "playing a chord" action, and need not decompose the observed action in its constituents (Vogt *et al*. 2007).

endorse, what the subjects in (Buccino *et al.* 2004a) recombined were actually basic sensory images. So, if the predictive processing account of mirror neurons is correct, it appears that basic images can be recombined as (4) requires.

Lastly, what about (5)? Are these basic sensory images decouplable? Again, there are reasons to answer positively. A substantial body of evidence (e.g. Ganis, Thompson and Kosslyn 2004; Alkhadi *et al.* 2005) speaks of the fact that in both visual and motor imagery cortical regions subserving on-line action and perception are active, sometimes including also the primary visual and motor cortices (e.g. Miller *et al.* 2010; Albers *et al.* 2013). But if the predictive processing story radical enactivists adhere to is correct, these areas are mainly in the task of generating a downward flow of virtual sensory inputs. Given that in imagery tasks subjects are not coupled with the imagined object, and given that an object is imagined (presumably) due to the underlying neural activity, it then seems that basic sensory images can be produced even when their target is absent. Hence, basic sensory images are decouplable from their targets.

So, basic sensory images satisfy (1) to (5) in conjunction. But satisfying (1) to (5) in conjunction provides a sufficient reason to identify something as a vehicle of content. Thus we have a sufficient reason to identify basic images as vehicles of content. As a consequence, they cannot be contentless, contrary to what radical enactivism holds.

I wish to close this section trying to anticipate two possible objections. Radical enactivists might be tempted to claim that I am begging the question against their position, and in particular against the "hard problem of content" (see Hutto and Myin 2013: 57-82). The "hard problem of content" is the analysis radical enactivists provide to claim that some cognitive processes and phenomena involve no content. As I understand it, the "hard problem of content" is basically a pessimistic meta-induction on the prospect of providing a satisfactory account of content in reductive and naturalistic terms (see Hutto and Myin 2018: 106-112). The argument considers a number of naturalistic and reductionist accounts of

content, and shows that none of them can uniquely determine the conditions of satisfaction of a vehicle of content. Among the accounts considered in the argument, there are also teleo-informational accounts. But conditions (1) to (5) look exactly like the sort of conditions that constitute a teleo-informational account of content (see Rowlands 2018: 349). So, am I begging the question against the "hard problem of content"?

I do not think I am. Suppose that the analysis radical enactivists provide when formulating the "hard problem of content" is correct. What that analysis shows is that the set of conditions imposed by teleo-informational accounts cannot *determine* the content of a vehicle. In other words, it shows that the set of conditions put forth by teleo-informational accounts cannot be what in virtue of which a vehicle acquires its content. But the claim I'm making here is *not* that basic imaginings acquire content in virtue of the fact that they satisfy (1) to (5). I am only claiming that we can identify them as vehicles of content because they satisfy (1) to (5). My claim is an epistemic claim on how to identify vehicles of content, not a metaphysical claim about what determines their content.[23]

An analogy might help clarify this point. Suppose you know Mr. X has been incarcerated, and that you do not have any reason to doubt his trial was fair and that it delivered a correct sentence. Now you have a sufficient reason to hold that Mr. X is a criminal. But you still might not know what made Mr. X a criminal. You can know that Mr. X is a criminal without knowing in virtue of what he is a criminal (i.e. what felony he committed). Similarly, my argument, if correct, provides a sufficient reason to identify basic images as vehicles of content, without pointing to the facts or the conditions that make them contentful. I'm just leveraging the dissociation between representational function and content determination Ramsey's framework affords. My argument, to function, does not need to presuppose that (1) to (5) are jointly sufficient to determine the mental content of an item. Therefore, my argument does not beg the question against the "hard problem of content".

[23] In other words, I'm *not* suggesting that conditions (1) to (5) are what "makes" something a vehicle of content. I'm just claiming (1) to (5) allow us to recognize a vehicle of content as such.

Radical enactivists might also claim that only the culturally regimented activity can *possibly* qualify as contentful (Myin 2020; Myin and van den Herick 2020). But surely the production of basic sensory images is not the activity of an agent. In fact, it is the activity of an agent's part; namely the agent's brain. Hence, it cannot possibly qualify as content-involving.[24]

Thus articulated, the argument seems question begging (to me): after all, the point I am making questions the assertion that only the socially regimented activity of enculturated agents can be contentful. Hence, radical enactivists need some independent reason to support their position. If I understand radical enactivists correctly, their position is supported by the "hard problem of content". However, the "hard problem of content" does not *show* that only certain activities of enculturated agents can be contentful. The "hard problem of content" shows only that some theories of content run into terrible problems when it comes to determine the intensional (with-an-"s") aspect of mental content. But, as it is presented in (Hutto and Myin 2013: 57-82; 2018: 106-111), the "hard problem of content" does not consider a number of theories of content, such as conceptual role semantics (e.g. Block 1998) or more modern informational theories that do not ground semantic content in covariance relations (see Weissglass 2019 for one such theory that can be straightforwardly applied to predictive processing).[25] Being silent on these, the "hard problem of content" does not show that all naturalistic theories of content fail (Shapiro 2014). Thus, it does not support the claim that content can be found *exclusively* within the socially regimented activities of enculturated agents. Hence, there is little reason to believe that only the socially regimented activities of enculturated agents can be contentful.

[24] Thanks to the anonymous reviewers for having urged me to make this point more explicit.

[25] Notice: I'm not claiming that these theories successfully naturalize content. I'm just claiming that radical enactivists have not shown that these theories fail too. And given that it is the radical enactivist the one asserting that the "hard problem of content" forces us to consider contentful only the activities of enculturated agents, it is the radical enactivist the one who bears the burden of proof.

5. Concluding remarks

In this essay, I've provided a sufficient reason to consider basic sensory images as vehicles of content, so as to counter the radical enactivist's claim that we should conceive them as contentless. One might wonder what my argument is good for. To what end prove that basic images are contentful, if one does not have a good (naturalistic and reductionistic) account of content? The answer is: to block the way radical enactivists deal with the "hard problem of content". As I understand them, their favorite strategy is just that of *walking away* from that problem, declaring that most of cognitive activities are contentless (see Hutto and Myin 2018: 105). But if the argument I presented here is correct, now there is a theoretical reason[26] not to allow them to walk away from the problem. If my reasoning is on the right track, there is a theoretic, sufficient reason to consider these basic images (alongside the perceptuomotor interactions they animate) as contentful. So, just as anyone else, radical enactivists must face the hard problem of content. They are not allowed to just walk away from it.

References

Adams, R., Perrinet, L. U., Friston, K. (2012). Smooth pursuit and visual occlusion: active inference and oculomotor control in schizophrenia. *PLoS One*, *7*(10), e:47502.

Adams, R., Shipp, S., Friston, K. (2013). Predictions, not commands: active inference in the motor system. *Brain Structure and Function*, *218*(3), 611-643.

Adams, R., *et al.* (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, *4*:47.

[26] Aside from pre-theoretical ones. Intuitively, we would describe our mental states as having conditions of satisfaction. We intuitively hold that our perceptions can be accurate, for instance.

Albers, A. M. *et al.* (2013). Shared representation for working memory and mental imagery in early visual cortex. *Current Biology 23*(15), 1427-1431.

Alkhadi, H., *et al.* (2005). What disconnection tells about motor imagery: evidence from paraplegic patients. *Cerebral Cortex*, *15*(2), 131-140.

Ambrosini, E., Costantini, M., Sinigaglia, C. (2011). Grasping with the eyes. *Journal of Neuropsychology*, *106*(3), 1437-1442.

Ambrosini, E. *et al.* (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS one*, *8*(7): e67916

Arcangeli, M. (2019). The two faces of mental imagery. *Philosophy and Phenomenological Research*, *101*(2), 304-322.

Block, N. (1998). Conceptual role semantics. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*, (pp. 242-256). New York: Routledge.

Buccino, G., *et al.* (2004a). Neural circuits underlying the imitation of hand actions: an event related fMRI study. *Neuron*, *42*(2), 323-334.

Buccino, G. *et al.* (2004b). Neural circuits involved in the recognition of actions performed by non-conspecifics: an fMRI study. *Journal of Cognitive Neuroscience*, *16*(1), 114-126.

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

Clark, A. (2012). Dreaming the whole cat: generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind*, *121*(483), 753-771.

Clark, A. (2014). Perceiving as predicting. In D. Strokes, M Matthen, S. Biggs (Eds.), *Perception and its Modalities* (pp. 23-43). New York: Oxford University Press.

Clark, A. (2016). *Surfing Uncertainty*. New York: Oxford University Press.

Clark, A. (2018). Beyond the "Bayesian blur": predictive processing and the nature of subjective experience. *Journal of Consciousness Studies*, *25*(3-4), 71-87.

Costantini, M., Ambrosini, E., Sinigaglia, C. (2012). Out of your hand's reach, out of my eye's reach. *Quarterly Journal of Experimental Psychology*, *65*(5), 848-855.

Dołega, K. (2017). Moderate predictive processing. In T. Metzinger, W. Wiese (Eds.), *Philosophy and Predictive Processing*: 10. Frankfurt am Main, the MIND Group. https://doi.org/10. 15502/9783958573116

Downey, A. (2018). Predictive processing and the representation wars: a victory for the eliminativist (via fictionalism). *Synthese*, *195*(12), 5115-5139.

Dretske, F. (1988). *Explaining Behavior*. Cambridge, MA.: The MIT Press.

FitzGerald, T., Dolan, R., & Friston, K. (2014). Model averaging, optimal inference, and habit formation. *Frontiers in Human Neuroscience*, *8*:457.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815-836.

Friston, K., *et al.* (2016). Active inference and learning. *Neuroscience and Behavioral Reviews*, *68*, 862-879.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1-2), 137-160.

Ganis, G., Thompson, W. L., & Kosslyn, S. (2004). Brain areas underlying visual mental imagery and visual perception: a fMRI study. *Cognitive Brain Research*, *20*(2), 226-241.

Gładziejewski, P. (2015). Explaining cognitive phenomena with internal representations: a mechanistic perspective. *Studies in Logic, Grammar and Rhetoric*, *40*(1), 63-90.

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, *193*(2), 559-582.

Haugeland, J. (1991). Representational genera. In W. Ramsey, S. P. Stich, D. E. Rumelhart (Eds.), *Philosophy and Connectionist Theory* (pp. 91-114). New York: Rutledge.

Hohwy, J. (2013). *The Predictive Mind*. New York: Oxford University Press.

Hurley, S. (1998). *Consciousness in Action*. Cambridge, MA.: Cambridge University Press.

Hutto, D. (2013). Why believe in contentless beliefs?. In Nottelmann, N. (Ed.), *New Essays on Belief* (pp. 55-74). London: Palgrave Macmillan.

Hutto, D. (2015). Overly enactive imagination? Radically re-imagining imagining. *The Southern Journal of Philosophy*, *53*, 68-89.

Hutto, D. (2018). Getting into predictive processing's great guessing game: bootstrap heaven or hell?. *Synthese*, *195*(6), 2445-2458.

Hutto, D., Kirchhoff, M.; & Myin, E. (2014). Extensive enactivism: why keep it all in?. *Frontiers in Human Neuroscience*, *8*: 706.

Hutto, D., & Myin, E. (2013). *Radicalizing Enactivism*. Cambridge, MA.: The MIT Press.

Hutto, D., & Myin, E. (2017). *Evolving Enactivism*. Cambridge, MA.: The MIT Press.

Hutto, D., & Myin, E. (2018). Going radical. In A. Newen, L. De Bruin, S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition* (pp. 95-116). New York: Oxford University Press.

Hutto, D., & Satne, G. (2015). The natural origins of content. *Philosophia*, *43*(3), 521-536.

Hutto, D., & Satne, G. (2017). Continuity skepticism in doubt: a radically enactive take. In C Durt, T. Fuchs, C. Tewes (Eds.), *Embodiment, Enaction and Culture* (pp. 107-129). Cambridge, MA.: The MIT Press.

Kilner, J. M., Friston ,K., & Frith, C. D. (2007). Predictive coding: an account of the mirror neurons system. *Cognitive Processing*, *8*(3), 159-166.

Kirchhoff, M. D., & Robertson, I. (2018). Enactivism and predictive processing: a non-representational view. *Philosophical Explorations*, *21*(2), 264-281.

Langland-Hassan, P. (2015). Imaginative attitudes. *Philosophy and Phenomenological Research*, *90*(3), 664-686.

Menary, R. (2007). *Cognitive Integration - Mind and Cognition Unbounded*. New York: Palagrave MacMillan.

Mesulam, M. (2008). Representation, inference, and trascendent encoding in neurocognitive networks of the human brain. *Annals of Neurology*, *64*(4), 367-378.

Miller, K. J., *et al*. (2010). Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, *107*(9), 4430-4435.

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. Cambridge, MA.: The MIT Press.

Myin, E. (2020). On the importance of correctly locating content: why and how REC can afford affordance perception. *Synthese*, https://doi.org/10.1007/s11229-020-02607-1

Myin, E., & van den Herik, C. A. (2020). A twofold tale of one mind: revisiting REC's multi-storey story. *Synthese*, https://doi.org/10.1007/s11229-020-02857-z

Orlandi, N. (2014). *The Innocent Eye*. New York: Oxford University Press.

Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*(9), 451-456.

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

Ramsey, W. (2016). Untangling two questions about mental representation. *New Ideas in Psychology*, *40*, 3-12.

Rizzolatti, G., & Sinigaglia, C. (2006). *Mirrors in the Brain*. New York: Oxford University Press.

Rowlands, M. (2006). *Body Language*. Cambridge, MA.: The MIT Press.

Rowlands, M. (2018). Disclosing the world: intentionality and 4E cognition. In A. Newen, L. De Bruin, S. Gallagher (Eds.), *The Oxford Handbook of 4E Cognition* (pp. 335-353). New York: Oxford University Press.

Shapiro, L. A. (2014). Radicalizing enactivism: basic minds without content, by Daniel D. Hutto and Erick Myin. *Mind*, *123*(489), 213-220.

Shea, N. (2018). *Representation in Cognitive Science*. New York: Oxford University Press.

Tani, J. (2016). *Exploring Robotic Minds*. New York: Oxford University Press.

Thompson, E. (2007). *Mind in Life*. Cambridge, MA.: Harvard University Press.

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

Vogt, S. *et al*. (2007). Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage*, *37*(4), 1371-1383.

Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: the varieties of enactivism. *Topoi*, *36*(3), 365-375.

Weissglass, D. E. (2019). Greatest surprise reduction semantics: an information theoretic solution to misrepresentation and disjunction. *Philosophical Studies*, *177*, 2185-2205.

Wiltsher, N. (2019). Characterizing the imaginative attitude. *Philosophical Papers*, *48*(3), 437-469.