Enactivism and the 'Explanatory Trap'. A Wittgensteinian Perspective

ANNA BONCOMPAGNI

Dipartimento di Filosofia,
Comunicazione e Spettacolo.
Università di Roma Tre
anna.boncompagni@uniroma3.it

Abstract

This paper explores the enactive approach in cognitive science with an eye on the later Wittgenstein's philosophy. The aim is not that of answering the question: was Wittgenstein an *ante litteram* enactivist? He was not, because he was not an *ante litteram* (cognitive) scientist of any kind. The aim, conversely, is that of answering the question: can enactivism be Wittgensteinian? In answering positively, it will be argued that a Wittgensteinian framework can help enactive cognitive scientists in dissolving certain old problems which they sometimes seem not to be able to get rid of. After the Introduction, the first two sections of the paper concern the Wittgensteinian standpoint on psychological concepts (Section 2) and the enactivist approach in its general terms (Section 3). Section 4 attempts a closer examination of some key concepts – chiefly representations, the inner, the "explanatory gap", the "hard problem" of consciousness – considering both the enactivists' and Wittgenstein's attitude towards them. The Conclusion surmises the benefits of a Wittgensteinian perspective also hinting at some other problems which it can help to clarify.

1 Introduction

Enactivism, initially proposed at the beginning of the Nineties, is nowadays one of the most relevant approaches in the field of the cognitive sciences. Although thirty years have already passed since it appeared, there is still not much clarity about its precise boundaries and its relation to other similar perspectives, such as ecological psychology, embodied cognition, sensorimotor contingencies account, neurophenomenology, extended mind cognition and the like, usually and confusedly grouped together under the label of the "4E" cognition approaches, "E" standing for embeddedness, extendedness, embodied, enactive (Menary 2010). A unifying factor of these accounts is their common opposition to traditional cognitivism, the latter usually being described as the study of cognition seen as a mental computation based on inner representations. The philosophical ancestors typically singled out by enactivists are Edmond Husserl, Maurice Merleau-Ponty, Martin Heidegger, John Dewey, and sometimes William James and Ludwig Wittgenstein. My aim in this paper is to explore the philosophy of cognitive sciences by focusing on *enactivism* and on *Wittgenstein*. The two perspectives seem indeed to share a common view on some key issues of cognitive sciences, but a further inquiry is needed in order to bear out whether and to what extent this is true.

I do not mean to suggest that Wittgenstein can be read as an enactivist ante litteram: Wittgenstein was not a scientist and always considered the methods of science as extremely misleading for philosophers[1]. Inasmuch as enactivism is a scientific perspective, Wittgenstein would not be an enactivist, because he would not be a (cognitive) scientist of any kind. What I am interested in suggesting, however, is that a Wittgensteinian framework for cognitive science is a perspective worth working on. This is especially the case for enactivists, on the one hand because the angle from which they conceive the key notions of cognitive science fits naturally into Wittgenstein's viewpoint, on the other hand because there are some persistent problems against which they struggle, which could be clarified through a Wittgensteinian, rather than a phenomenologist, strategy. I am referring here mainly to two connected themes: the so-called "explanatory gap" (Levine 1983) concerning how a first person perspective could gain a respectable third person scientific status; and the "hard problem" of consciousness, according to which the explanatory gap just mentioned can only be filled through a scientific account of how the brain can generate phenomenal experience (Chalmers 1995). At the core of both aspects, which I will characterize in more detail in the following pages, is the notion of qualia, the "what it is like" character of subjective experiences.

The paper proceeds as follows. Section 2 offers an outlook of Wittgenstein's philosophy of psychology as well as of his assessment about the relation between

---

[1] See also Wittgenstein (1969a).

philosophy and science. Maxwell Bennett and Peter M. S. Hacker's work on the philosophical foundations of neuroscience will also be introduced[2]. Section 3 presents the main features of enactivism, focusing on the original formulation of the project by Francisco Varela, Eric Thompson and Eleanor Rosch[3], on the sensorimotor account by Kevin O'Regan and Alva Noë[4], on the project of a neurophenomenology outlined by Varela and developed also by others[5], and on Daniel D. Hutto and Erik Myin's radical enactivism[6] (other more or less akin perspectives will not be considered, for obvious limitations of space). The following section deals with some key concepts on which enactivists debate – representations, the inner, first person and third person accounts, the "explanatory gap", the "hard problem" of consciousness – and proposes a Wittgensteinian reading of them. It will be argued that the "explanatory gap", because of the terms in which the matter is put, is actually an explanatory *trap*, which can be avoided through a different description of what is at stake. Consequently, the "hard problem", in spite of its apparent hardness, is a *false* problem. The Conclusion surmises the benefits of a Wittgensteinan framework and briefly hints at other issues which could be addressed.

2 A Wittgensteinian framework

In the last page of his *Philosophical Investigations*, Wittgenstein famously pointed out that the problems of psychology are not due to its being too young a science and in need of more empirical research, but to the mixture of "experimental methods and conceptual confusion" which characterizes it[7]. A remark, like many others in his philosophy of psychology, probably originating from his reading of William James[8]. And a remark which today could refer as well to cognitive sciences and neuroscience. The positive side of this critique is that it is possible to avoid some of the problems which psychology and cognitive sciences face, through a conceptual dissolution of their confusions, and a corresponding conceptual clarification of the main topics which they attempt to deal with. This is a philosophical task, and the fulfillment of this task has as its precondition a definite demarcation of

---

[2]Bennett and Hacker (2003)

[3]Varela et al. (1991)

[4]O'Regan and Noë (2001);Noë (2004).

[5]Varela (1996);Petitot et al. (1999).

[6]Hutto (2006b); Hutto and Myin (2013).

[7]Wittgenstein (2001, II, p.197).

[8]It is easy to read Wittgenstein's statement, according to which psychology's state "is not comparable with that of physics, for instance in its beginnings" (Wittgenstein 2001, II, p.197), as a direct comment on William James's words: "At present - he wrote in a text that Wittgenstein knew very well - psychology is in the condition of physics before Galileo and the laws of motion, of chemistry before Lavoisier and the notion that mass is preserved in all reactions" (James 1892, p.401). See also Boncompagni (2012a).

the respective fields and methods of science and philosophy. The demarcation, in Wittgenstein's perspective, is so deep as not to allow any overlapping between the two, because allowing that would amount to committing a categorical mistake: confusing description with explanation, grammatical with empirical, sense and nonsense with true and false, reasons with causes (Tripodi 2009). Philosophy is concerned with the former in each pair, science with the latter.

The warning that philosophy should not attempt to enter the field of science (but the converse also holds) leaves the possibility open in any case that a scientific work may benefit from the conceptual clarifications offered by a philosophical outlook. It is in this spirit, that a Wittgensteinian framework can constitute a fruitful starting point for cognitive sciences.

Psychological concepts are a pervasive topic in Wittgenstein's later philosophy, especially from the second half of the Forties until the last days of his life and work, as testified in the *Nachlass* and in published writings such as Part II of the *Philosophical Investigations*, the *Last Writings on the Philosophy of Psychology*, the *Last Writings on the Philosophy of Psychology, the Remarks on the Philosophy of Psychology, On Certainty* and in the collection of notes from his lectures on philosophical psychology, edited by Peter Geach[9]. In the Wittgensteinian perspective, the objects of cognitive sciences broadly construed, namely, perception, sensation, emotion, intuition, feeling, pain, desire, consciousness, belief, hope, understanding, memory, will, reason, imagination and so on, can have neither a mentalistic nor a physicalistic reductionist explanation. Starting from a linguistic and conceptual analysis, his purpose is to make explicit, via a focus on ordinary language practices, how these apparently private phenomena are actually public and overtly available, and how they are connected to actions, praxis, linguistic games, and ultimately to forms of life. The claim to the publicity of what is usually called "mental states", does not amount to a behaviouristic point of view: the existence of internal feelings and thoughts is not denied. It is the metaphysics of the inner/outer dichotomy which is contested. In order to clarify the concepts involved here, Wittgenstein makes use of different strategies, including the so-called "private language argument"[10], but also many subtle and detailed descriptions of ordinary language practices, as well as imaginative and sometimes "science-fiction like" explorations of the limits of sense and nonsense, that is, of the limits of our form of life. Mental phenomena are thus brought back to their environment, the circumstances in which they occur, and this leads to acknowledging the centrality of the whole life of a person and of the whole of her social interactions in order to understand mind itself. In a nutshell, at the core of a hypothetical Wittgensteinian cognitive science is the inextricable connection between mind and action, knowledge and practice, language and life,

---

[9]Respectively (Wittgenstein (2000); Wittgenstein (2001); Wittgenstein (1982); Wittgenstein (1992); Wittgenstein (1980); Wittgenstein (1969b); Wittgenstein (1988)).

[10]Wittgenstein (2001, §§243-315).

nature and culture.

It may be useful for our purposes - in addition to Wittgenstein's words, which we will directly employ in what follows - to adapt to contemporary cognitive science the meticulous analysis of the (often mistaken) conceptual bases of neurosciences, which a Wittgensteinan philosopher like Peter M. S. Hacker and a neuroscientist like Maxwell R. Bennett published ten years ago (Bennett and Hacker 2003). Although it endorses a rather traditional interpretation of Wittgenstein's philosophy and may at times turn out to be irritating in its insistence on the correctness of Wittgenstein's view, defending him against any kind of possible criticism or un-orthodox interpretation, Bennett and Hacker's work is interesting because it focuses precisely on the philosophical task of conceptual clarification, sharply distinguishing the philosophical and the scientific activities, but without disregarding either. In their (and partly Wittgenstein's) words, "what needs to be said can be said clearly, and saying it clearly will benefit, not diminish, the actual achievements of neuroscience"[11].

In tracing the origins of the philosophical misconceptions which accompany neuroscientific research, in this work a particular grudge is, not surprisingly, against the figure of Descartes, considered as the main responsible for the mind-body dualism usually characterizing scientific inquiries in the past centuries. This dualism has nowadays assumed a new garment: it often appears in the "degenerated form" of a brain-body dualism, where the brain has substituted the mind and is the subject of volitional psychological attributes [12]. There are actually more interconnected problems here. One is the classical question of how mind/brain and body are related and how they can have a causal power on each other. A second problem is the "explanatory gap": how can we bridge the gap between a third person objective description of the brain, and a first person subjective description of experience? A third related problem the so-called "hard problem" of consciousness, that is, how does the brain generate experience? The answer to this problem is generally deemed to be necessary in order to fill the "explanatory gap". And finally there is what Bennett and Hacker (2003, p.68 ff.) call the mereological fallacy of neuroscience, the fallacy of attributing to a part of an organism (usually the soul or the mind in the past, and the brain in the present) properties which belong to the whole organism or person (being conscious, having feelings or emotions, thinking, etc). One more confusion which Bennett and Hacker refer to more than once is the idea of mental inner representations, and this, again, is a concept strictly tied up with the dualistic Cartesian conception of the relation between mind and matter. These problems are deeply intertwined, implying and sustaining each other in a net of explicit and implicit cross-references.

---

[11]Bennett and Hacker (2003, p.107).
[12]Bennett and Hacker (2003, p.111).

A Wittgensteinian framework, contrasting and dissolving the inner-outer dichotomy, demands a radical shift in the traditional cognitivist and neuroscientific project of inquiring into the workings of the mind and/or brain, towards a perspective centered on the embodied, intersubjective, living, and at the same time cultural and linguistic character of human experience. Section 4 will develop these ideas, now only sketched, and try to make use of them in connection to the most debated issues in enactivism. But before we can do that, let me focus, in Section 3, on enactivism, trying to single out its main features.

3 Enactivism

The shift towards a perspective centered on embodiment and intersubjectivity, which a Wittgensteinian framework suggests, is precisely what is at the core of enactivism. Many philosophers and cognitive scientists today use this term in various ways and often non coincidently, and in order not to generate more confusion it is worthy to point out its main features by referring to the commonly accepted origin: *The Embodied Mind* (Varela et al. 1991). It is here that we find the first proposal of the term "enaction"[13], or enactive cognition [14]:

> We propose as a name the term enactive to emphasize the growing conviction that cognition is not representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs.

What enactivism aims at refuting, on the basis of this definition, is the idea of a mental representation as the necessary correlate of the "pregivenness" of the mind on one side, and the world on the other side. The "logical geography of inner versus outer"[15] which is a consequence of this picture is resolutely refused, or better said bypassed, through the definition of cognition as "embodied action", and the specification that

> [b]y using the term *embodied* we mean to enlight two points: first, that cognition depends upon the kinds of experience that come from having a body with various sensorimotor capacities, and second, that these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological, and cultural context. By using the term *action* we mean to emphasize once again that sensory and motor processes, perception and action, are fundamentally inseparable in lived cognition. (...) [T]he enactive approach consists of two

---

[13]Varela et al. (1991, p.9).

[14]For more recent accounts, see the introductory essay in Stewart (2010).

[15]Varela et al. (1991, p.172).

points: (1) perception consists in perceptually guided action and (2) cognitive structures emerge from the recurrent sensorimotor patterns that enable action to be perceptually guided[16].

Sensorimotor patterns, contingencies, capacities, laws is a key concept also emphasized by O'Regan and Noë's study of vision and visual consciousness (O'Regan and Noë 2001), as well as in Noë's subsequent study of perception (Noë 2004). The active role of the organism is the starting point for the idea that sensation and movement are inextricably interconnected in such a way that only if an organism can actively interact with its environment, can it be said to be conscious, to perceive and to possess cognition. Again, this claim is bound to a form of antirepresentationalism:

> Instead of assuming that vision consists in the creation of an internal representation of the outside world whose activation somehow generates visual experience, we propose to treat vision as an *exploratory activity*. The central idea of our new approach is that *vision is a mode of exploration of the world that is mediated by knowledge of what we call sensorimotor contingencies* [17].

These contingencies or patterns are the regularities which govern the relation between perceiver and object. It is thanks to a practical knowledge of these regularities that we ordinarily perceive: roughly, we perceive an apple as an apple, and not as a red spot, because we implicitly know that by moving ourselves to one side other aspects of the apple will appear. This knowledge, regarding both sensations and real or potential movements, is what enables us to perceive the world as it appears, and, through its appearances, as it is[18].

Noë (2004) conclusion partly differs from O'Regan and Noë (2001). This is interesting, because the difference concerns the problem of consciousness. In the 2001 paper consciousness was not a direct object of inquiry, and the authors spoke about it only in terms of a description of sensorimotor knowledge. In his 2004 re-reading of that strategy, Noë affirmed that it "purchased noncircularity and explanatory power at the expense of living up phenomenological aptness" [19]. In other words, it lacked a true phenomenological account of experience:

> A creature enjoys phenomenally conscious perceptual states when it
> has knowledge of the relevant patterns of dependence of neural activity

---

[16]Varela et al. (1991, p.173).

[17]O'Regan and Noë (2001, p.940).

[18]See also Noë (2004, ch.3,5). Noë is also committed to a conceptual theory of perception, derived from Kant and McDowell, although he specifies that sensorimotor abilities are only "protoconceptual", that knowledge is practical and not theoretical and that it is a sort of know-how. More on this shortly.

[19]Noë (2004, pp.228-229).

on movement. But how can phenomenally unconscious states of this sort be the basis of phenomenal consciousness? This question remains unanswered. [20]

The unanswered question is a version of the "explanatory gap", and whereas O'Regan and Noë (2001) affirmed that "qualia are an illusion, and the explanatory gap is no real gap at all"[21], Noë (2004) conclusion is an acknowledgment that the explanatory gap is still there and the sensorimotor approach has failed in solving it. According to him, it is probably only through an evolutionary perspective centered on the notion of life, and on the question of how life has originated, that this problem can be correctly faced[22].

The evolutionary perspective and the connected notion of life as a promising path in the direction of the filling of the gap is explicitly endorsed by more recent enactivist works, including Thompson's follow-up of *The Embodied Mind* (Thompson 2007)[23]. Thompson affirms that enactivism can offer important resources for making progress on this problem, and his work wants to constitute a contribution towards what has been called "neurophenomenology", that is, the project of a naturalization of phenomenology [24].

At first sight, even someone only superficially familiar with phenomenology would be perplexed in the face of a perspective bearing the name of "neurophenomenology". Husserl himself worked exactly the other way around, trying to detach ourselves from a natural (or naturalizing) attitude, towards a phenomenological analysis of how it can develop. To fully understand what is implicated in this idea, we need to turn to its original formulation, once again originating in Varela.

Varela (1996) clarifies that the aim of neurophenomenology has nothing to do with a scientist project of naturalization aiming at transforming philosophy into a natural science. Quite the opposite, the project plans an entirely new framework, within which the "reciprocal constraints" [25] are investigated wherein the structures of phenomenological experience and their counterparts in cognitive sciences are related[26]. The fundamental idea is that conscious experience can be studied through phenomenology, by exercising the method of phenomenological reduction. Experience, said Varela, "demands a specific examination in order to free it from its status as habitual belief"[27], and this examination can be conducted through reduc-

---

[20]Noë (2004, p.229)

[21]O'Regan and Noë (2001, p.960).

[22]But the doubt still remains whether in this way we are only "substituting one explanatory gap for another" (Noë 2004, p. 231).

[23]See also Stewart (2010).

[24]Thompson (2007, p.14).

[25]Varela (1996, p.343).

[26]See also Varela (1996, p.330) and Gallagher (2010).

[27]Varela (1996, p.336).

tion. Phenomenological reduction is at odds with physical reductionism. It is akin to doubt and can be described as

> a sudden, transient suspension of beliefs about what is being examined, a putting in abeyance our habitual discourse about something, a bracketing the pre-set structuring that constitutes the ubiquitous background of everyday life[28].

At the basis of this attitude is the conviction that reduction is possible and that it gives us access to experience how it actually is, apart from our prejudices. This is the source of what Varela interestingly calls a "remedy" for the hard problem [29]. It must be said that Varela's methodological stance and its difference from a scientific idea of naturalization is not always clear in literature, and under the label of the naturalization of phenomenology stands a variety of approaches sometimes hard to reconcile with each other (Petitot et al. 1999). Thompson's own idea of continuity between life and mind, for example, entails a continuity between science and phenomenology which, as he acknowledges, challenges Husserl's perspective [30], and maybe Varela's as well.

The last enactivist approach which I would like to mention is Daniel D. Hutto and Erik Myin's radical enactivism (Hutto (2006b); Hutto and Myin (2013)). Although it shares the general frame with other kinds of enactivism, focusing on the embodied organism as the only method by which it is possible to account for cognition, it aims at being 'radical' in defending a strong antirepresentationalism and in considering the basic forms of emotional experience and response as non conceptual and non contentful. Basic minds[31] are intentionally directed towards aspects of the environment, but this kind of intentionality does not need to be explained in terms of content or concept. Working on Gilbert Ryle's distinction between knowing-how and knowing-that, and on his insistence on the fact that "knowing-how is not reducible to any sandwich of knowings-that" [32], Hutto criticizes the sensorimotor enactivism for its misuse of the concept of knowledge[33] and for its reliance on conceptuality.[34] Although O'Regan and Noë (2001, pp.944, 946) and Noë (2004, pp.11, 88, 118) repeatedly insist that the kind of knowledge involved in their approach is practical knowledge, that is, know-how (the two notions may not equate, but they seem to conflate them), and although Noë (2004, p.210) opposes Stanley and Williamson (2001) attempt to demonstrate that knowledge-how is just a species of knowledge-that, they in fact speak, on many occasions, of know-

---

[28]Varela (1996, p.337).

[29]See section 4.

[30]Thompson (2007, p.356).

[31]Hutto and Myin (2013, pp.11-13).

[32]Ryle (1945, p.15).

[33]The same criticism can be found in Rowlands (2007, pp.435-436).

[34]Hutto (2005); (Hutto and Myin 2013, p. 25–ff.).

ing *that*. Typically, our expectations about perception are considered as a form of knowing that by moving the eyes towards that direction, that aspect of an object would appear[35]. Moreover, in dealing with the activities of the brain, they use verbs such as "to judge", "to assume", "to conclude"[36], thereby also committing what Bennett and Hacker (2003) call the mereological fallacy[37]. Hutto and Myin (2013, p.34 ff.) also criticize the autopoietic account – that is, Varela et al. (1991) and Thompson (2007) – for its making too much use of notions like interpretation, meaning, sense-making, understanding. These are forms of conservative enactivism, he holds, while enactivism needs to be radicalized in order to constitute a serious and consistent alternative to cognitivism. But the proponents of radical enactivism are careful in specifying that they are not denying that mind and thought *can be* - and sometimes undoubtedly *are* - characterized by representational contents and concepts: "What is ruled out by REC [Radical Enactivist Cognition] is that content-involving mentality is *basic* and is found in *any and all forms of mentality*"[38].

So far, we have examined only a small portion of the existent literature on enaction and we can already see the emergence of some key-problems on which there is not a complete agreement. The next section is devoted to a discussion of representations, the "explanatory gap" and the "hard problem" of consciousness, and is meant to show how a wittgensteinian perspective on them could be of help. As mentioned, Wittgenstein is rarely cited as one source of reflection for enactivists: philosophical debts are more often declared towards the phenomenologist tradition (Husserl, Heidegger, Merleau-Ponty, sometimes Sartre), and in some cases towards pragmatism (particularly Dewey, sometimes James, rarely the neopragmatists) and Eastern thought (the comparison between an embodied theory of the mind and the Buddhist tradition of Madhyamika is central in Varela et al. (1991)). As we shall see, the prevalence of a phenomenological framework is actually one of the sources of some of the permanent problems of enactivism.

---

[35]O'Regan and Noë (2001, p.961,963) and Noë (2004, p.63,88).

[36]((O'Regan and Noë 2001, pp.949-951); (Hutto 2005, p.392); (Hutto 2006b, pp.23-24)).

[37]Bennett and Hacker (2003, p.68).

[38]Hutto and Myin (2013, p.13, my emphasis). It must be said that when Hutto and his adversaries speak about conceptuality and content, they do not seem to speak about the same subject. Noë's reasoning about experience echoes Kant and John McDowell, while the Kantian attitude is completely absent in Hutto's account. Hutto and Myin (2013, p.31) "find the notion of concept in play in Noë's account to be much too liberal and too individualistic", and think that in Noë "[t]he bar for being a concept user is set very low". Conversely, it could be said that for Hutto and Myin the bar is set very high. See for instance their definition of content,Hutto and Myin (2013, p.x,67). In any case, the matter is so complex and controversial and the space here so limited, that I must leave it aside. I will only hint at it in the concluding remarks.

4   Gaps and traps

One concept generally opposed by almost every (but not every) enactivist theorist is the concept of representation. As we have seen, enactivism itself was born in order to oppose cognitivism, and, as Varela et al. (1991) put it, "cognitivism is *mental representation*: the mind is thought to operate by manipulating symbols that represent features of the world or represent the world as being a certain way"[39].

What is denied is, chiefly, the idea of mental inner representations as a necessary (let alone sufficient) condition for cognition. Two themes are interconnected here: the conception of the mind as an inner realm; and the notion of representation. Joined together, these two themes amount to shape cognition as the activity of mirroring the external world, through representations, inside the inner realm of the mind. This Cartesian fashioned picture is still widespread both in philosophy and in cognitive science, as well as in common sense. Why is enactivism so determined in opposing it?

It must be said that enactivists are not affirming that representation does not play any role. They do acknowledge the relevance of representation for certain kinds of cognition[40]. But they deny that the *paradigm* of cognition should be based on representation. Even activities like perception or imagination need not be accounted for by using the concept of representation. Perception is indeed explained in terms of dynamical interactions and exploratory activities based on sensorimotor contingencies, and imagination in terms of a re-presentation (with the hyphen!) of perception. As Thompson (2007) elegantly surmises,

> a phenomenal mental image is not a phenomenal picture in the mind's eye, nor indeed is it any kind of static image or depiction; it is, rather, the mental activity of re-presenting an object by mentally evoking and subjectively simulating a perceptual experience of that object[41].

There is another, more controversial, aspect, in the enactivist battle against representation: they generally doubt the tenability of the hypothesis that an organism or a system in its basic behaviour acts on the basis of internal representations. According to some exponents, in fact, enactivism is compatible with "action-oriented representations" (Clark (1997); Mandik (2005)). Cappuccio and Wheeler (2012) for example underline that a bodily skilled coping with the environment can involve action-oriented representation: this is the case of sportsmen and sportswomen, actors, dancers who habitually optimize their performances by mentally representing them before execution. Moreover, according to Coates (2007), one of the central claims of enactivism, namely, the idea that perception

---

[39] Varela et al. (1991, p.8). See also O'Regan and Noë (2001, p.940) and Noë (2004, p.2).
[40] Varela et al. (1991, p.134), Noë (2004, pp.22-23) and Hutto and Myin (2013, p.13).
[41] Thompson (2007, p.297).

is essentially integrated with action, even *implies* the presence of inner representations. This is because the perception involved in what he calls "navigational activity", which according to him covers a large part of everyday life, requires the existence of inner representation [42]. The antirepresentationalist reply is that, in both cases, we are actually presented with peculiar, and not typical, activities. Even if we grant that in these cases representations are involved, pre-representation of performances and navigational activity are *not* the basis, nor the ordinary case, for action. Under general circumstances, actions are not pre-represented this way: we just act, without having to imagine in advance what we are to do or the goal we have to achieve[43].

A more decisive philosophical objection to the picture of mental representation is nuanced in Noë (2004) and in Hutto and Myin (2013), and can be traced back to Wittgenstein. Mentioning the Viennese philosopher, Noë writes: "anything that a picture in the head could do, could be done by a picture in the hand" [44]. And Hutto and Myin (2013) , while discussing the hypothesis of motor plans conveying instructive orders from the brain to the hand: "The trouble is that, even if we imagine that such representational contents exist, it is difficult to see how they could do the required work"[45]. This, I think, is a key point. Even if we suspend our judgment on the tenability of the representational hypothesis, and accept that there could be mental images in the mind or in the brain, whatever this could mean, the problem would remain untouched. Let us follow Wittgenstein's reasoning.

> If I give someone the order "fetch me a red flower from that meadow", how is he to know what sort of flower to bring, as I have only given him a *word*?
>
> Now the answer one might suggest first is that he went to look for a red flower carrying a red image in his mind, and comparing it with the flowers to see which of them had the colour of the image. Now there is such a way of searching, and it is not at all essential that the image we use should be a mental one. In fact the process may be this: I carry a chart co-ordinating names and coloured squares. When I hear the order "fetch me etc." I draw my finger across the chart from the word "red" to a certain square, and I go and look for a flower which has the

---

[42] Coates (2007, p.460).

[43] An interesting reformulation of the problem is proposed by Engel (2010). The criterion for success of cognitive operations, Engel points out, is not a 'veridical representation' of environmental features, but "viable action in a certain situation". For this reason, in developing what he calls a 'pragmatic turn' in cognitive sciences, he proposes the notion of 'directives', instead of that of action-oriented representations. For an alternative concept of action in enactivism, see also Rowlands (2007).

[44] Noë (2004, p.219).

[45] Hutto and Myin (2013, p.48)

same colour as the square[46].

So far, this is Noë's argument. The reasoning proceeds

> But this is not the only way of searching and it isn't the usual way. We go, look about us, walk up to a flower and pick it, without comparing it to anything. To see that the process of obeying the order can be of this kind, consider the order "*imagine* a red patch". You are not tempted in this case to think that *before* obeying you must have imagined a red patch to serve you as a pattern for the red patch which you were ordered to imagine[47].

Mental images as an instrument to explain how we act, or obey an order, or follow a rule, are useless. More: there is a flaw here, as the instruction thus conceived entails an infinite regress[48].The paradox is not a new one. Even supposing that I hold an image of a tree in my head, what is it in virtue of which this image is connected to the tree "out there"? Is it similarity? And how do we decide if the image is similar to the real object: is there another image, similar to both, connecting the two? It is not difficult to see here a new version of the familiar "third man argument", exploited by Aristotle against Plato.

Bennett and Hacker (2003), in their Wittgensteinian critique of the conceptual basis of neuroscience, deal at length with mental or inner representations. It is a Cartesian, together with a Lockean, inheritance, that is still at work in the background of neuroscientific and cognitive science research, providing the scheme according to which "to know" means to hold an image in the head[49]. But "to perceive is not to represent anything"[50], they state , and in any case it is a conceptual mistake to conceive of mental states or events as occurring in a part of the body of a person, namely, the brain, and not as concerning the person as a whole organism [51]. In perfect attunement with the enactivist idea that perception is direct and it is the perception of objects, not of images, they write:

> It is a mistake to suppose that what we perceive is always or even commonly an image, or that to perceive an object is to *have* an image of the object perceived. One does not perceive images or representations of objects, unless one perceives paintings or photographs of objects [52]
> .

---

[46]Wittgenstein (1969a, p.3).

[47]Wittgenstein (1969a, p.3).

[48]See also Boncompagni (2012b, p.33).

[49]A well-known philosophical criticism of this picture, its origins and consequences, is provided by Rorty (1979). See also Fischer (2011).

[50]Bennett and Hacker (2003, p.147).

[51]Bennett and Hacker (2003, p.112).

[52]Bennett and Hacker (2003, p.138, see also pp.192-193).

The enactivists' rejection of inner representations finds in a Wittgensteinian conceptual clarification the perfect ally, all the more so when this rejection is radical, not allowing the misunderstandings which typically lead to the arousal of difficulties and puzzles[53].

One even deeper problem that the concept of mental representations entails, used as a paradigm of cognition, is connected to the idea of the mind as an inner realm. Wittgenstein famously argued against the conception of sensations, emotions, thoughts as something essentially private, belonging to a separate ontological world – the inner. Many of his best known and debated remarks are devoted to this topic. Let me cite a few of them[54]:

> (...) Suppose everyone had a box with something in it: we call it a "beetle". No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at *his* beetle.- Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing.- But suppose the word "beetle" had a use in these people's language? - If so it would not be used as the name of a thing. The thing in the box has no place in the language-game at all; not even as a *something*: for the box might even be empty. (...)

> [I]f we construe the grammar of the expression of sensation on the model of 'object and designation' the object drops out of consideration as irrelevant.

> I can know what someone else is thinking, not what I am thinking. It is correct to say "I know what you are thinking", and wrong to say "I know what I am thinking.".

> (A whole cloud of philosophy condensed into a drop of grammar.)

> And hence also 'obeying a rule' is a practice. And to *think* one is obeying a rule is not to obey a rule.

> Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it[55].

The first point Wittgenstein is making is that it has literally no sense to think of sensations or thoughts as they were objects inside the mind, to which only the subject has access; indeed, if that were the case, no names could meaningfully be connected to such objects. The object itself would have no part in the game. The

---

[53] A parallel position with respect to this topic is Hubert Dreyfus' antirepresentationalism; see Dreyfus (2002).

[54] For a more extensive discussion, see Boncompagni (2012b, part 2).

[55] Wittgenstein (2001, I §202,293; II p.189). I have mixed up the remarks for the sake of argument.

philosophical outcome of this remark is surprising, clear and devastating. The consequence of considering an object as essentially private, is that the object drops out of the game. *The consequence of mentalism, is behaviourism.* As Perissinotto (1991, p.157, n.42) efficaciously puts it, Wittgenstein was not a behaviourist, because he was not a mentalist.

Second point: since it has no sense to conceive of emotions or thoughts as inner objects, we do not have an epistemic relationship with them. We feel, experience, think them, but it is not correct to say that we know them. On the contrary, we know other people's feelings and thoughts, in the ordinary sense in which, if I see someone in pain, I know she is in pain, or if I hear someone saying that she thinks so and so, I know she thinks so and so. Again, the philosophical outcome is devastating: with one single move it casts doubt on an ingenuous conception of self consciousness, and shows the problem of other minds to be grounded in a misconception.

Third point: given that the meaning of words cannot be fixed by means of a private ostensive definition, meanings are inescapably public; given that public criteria are grammatically needed in order for a rule to be a rule, no private rules are possible, that is, no private languages are possible. Language cannot be private because, if it is a language, then it is public.

This does not amount to a denial of the inner, but to a clarification of the conceptual relations among adjectives such as private, inner, hidden and the like. Ordinary language can give us the appropriate outlook from which to investigate these concepts. "If I lie to him and he guesses it from my face and tells me so - do I still have the feeling that what is in me is in no way accessible to him and hidden?" asks Wittgenstein, and replies: "Don't I feel rather that he sees right through me? It is only in particular cases that the inner is hidden from me, and in those cases it is not hidden because it is the inner."[56]. The relation between emotions or sensations felt by a person and her bodily or linguistic behaviour takes the form of *expressions*, which are not equivalent to reports or descriptions. There is a characteristic indeterminacy linking emotion and expression, and the way we usually understand others' feelings is characterized by an "imponderable evidence"[57] that "has nothing to do with either unbridgeable ontological divides or epistemological defects, and everything with the enormous variety and flexibility of human life"[58].

If this is correct, any attempt to conceive emotions, sensations, thoughts as internal objects is grounded in a fatal misconception and leads to irresolvable puz-

---

[56]Wittgenstein (1992, p.33).

[57]Wittgenstein (2001, II p.194).

[58]I am quoting here ter Hark (2004), who brilliantly highlights the Jamesian (and Darwinian) influence on Wittgenstein on these concepts. On the same subject see also Schulte (1995). On Wittgenstein and James in general, Goodman (2002).

zles. This is not generally the case of enactivist cognitive science, which thinks about experiences as interactive and avoids the conception of inner objects or representations. But the private language argument also has another relevant consequence which may affect enactivist approaches too. Here we find one reason why a Wittgensteinian point of view provides a better framework than a phenomenological point of view.

Enactivists are not committed to a characterization of the phenomenal aspect of experiences in "*qualia*-terms" [59], which precisely falls under Wittgenstein's criticism. But enactivists – we have hinted at it in the previous section – *are* sometimes committed to the idea that *there is* an explanatory gap which needs to be bridged, and that it is meaningful and possible to give accurate first person descriptions of inner experiences and confront them with or even translate them into third person descriptions. The phenomenological project, in a sense, is the project of a science of the first person. This idea, as we have sketched, for some enactivists assumes the form of neurophenomenology.

Now, it would be wrong to say that Wittgenstein is an "anti-phenomenologist" through and through. In his work there is a constant attention to the phenomenal aspect of experience. This was even his main concern in what has been identified as a phenomenological phase of his activity, at the end of the Twenties and at the very beginning of the Thirties, of which we can find detailed evidence in the *Philosophical Remarks*. But the later Wittgenstein's attention toward phenomenological experiences is always (and necessarily) linked to a corresponding attention to the circumstances in which they occur, to the way they are expressed, to the way in which they are connected to the contexts of the linguistic game and to the whole background of a form of life. There is nothing akin to a phenomenological description of *essences* in this kind of investigation. This difference is sometimes overlooked, and this is what happens, for example, in Michel Bitbol's attempt to find congruencies between the project of neurophenomenology and Wittgenstein's philosophy. According to Bitbol (2002), indeed, Varela's plan is consistent with the Wittgensteinian attitude towards phenomenological experiences. Let me first summarize Bitbol's position and then explain why, in my opinion, it is not correct.

Bitbol retains that the Husserlian conception of description of essences – the basis of Varela's neurophenomenology - is in many respects similar to the Wittgensteinian investigation into the relations among expression (first person), empathy (second person) and description (third person). In particular, according to Bitbol, both Varela - following Husserl – and Wittgenstein are interested in establishing stable correlations among these three categories, but while Wittgenstein remains within the domain of everyday use of psychological concepts, Varela extends and develops this perspective in the neuroscientific domain. Although it is interest-

[59]See also O'Regan and Noë (2001, p.961) .

ing to look at Wittgenstein's investigation from this standpoint, there are at least two reasons why it is not accurate to see the two perspectives as consistent with one another. The first reason is that the Wittgensteinian search for a perspicuous presentation of our psychological concepts *has to* remain within the domain of everyday experience, because this is precisely its spirit and objective: it aims at providing us with a clearer view of how we use our words and what we mean by using them, helping us to get rid of the pictures and mental cramps which hold us captives. Any step forward, or backward[60], in the direction of science and explanation misses the point, because trying to attend to our experiences in order to offer a satisfactory scientific description of them is precisely the first false move which may jeopardize the whole enterprise and its therapeutic aims:

> Philosophers constantly see the method of science before their eyes, and are irresistibly tempted to ask and answer questions in the way science does. This tendency is the real source of metaphysics, and leads the philosopher into complete darkness. I want to say here that it can never be our job to reduce anything to anything, or to explain anything[61].

It is not any kind of science, which can constitute a "remedy", to use Varela's expression, for our problems, especially if our problems are deemed to be due to the explanatory gap. The second, and connected, reason why the neurophenomenological project is not consistent with a Wittgensteinian framework is that the difference between psychological concepts and physical concepts is *not* the difference between first person and third person (Hutto 2006a). This is exactly what creates the gap, and cannot be what bridges the gap. What we have to attend to, when trying to elucidate the use of our words and psychological concepts, is not anything hidden which can be seen only from our first person point of view, but the contexts of actions and practices in which these concepts and words appear. And this is a direction that enactivism is in the position to recognize and to follow. Thus, if Wittgenstein can be said to be interesting and useful for cognitive science (Cole 2007), it is for his warning *against* the perspective of neurophenomenology, and not for his alleged convergence with it. This is especially true for some interpretations of neurophenomenology which do not take Varela's own methodological admonitions too seriously (Petitot et al. (1999); Thompson (2007)). The main point, let me repeat, is that the very ideas of an "explanatory gap" to be filled and of a "hard problem" of consciousness to be resolved depend on philosophical misconceptions. To put it as a slogan, the "explanatory gap" is an explanatory *trap*; and the "hard problem", in spite of its hardness, is a *false* problem.

---

[60]"It is so difficult to find the beginning. Or, better: it is difficult to begin at the beginning. And not try to go further back" Wittgenstein (1969b, §471).

[61]Wittgenstein (1969a, p.18).

Some enactivists' persistent attempts to move in the direction of a solution of the problem and of the filling of the gap, unfortunately, fail to notice that the main tenets of enaction itself point in the opposite direction, towards the dissolution of the problems. On the contrary, Hutto's Wittgensteinian background prevents him from engaging in the "mission impossible": indeed, when he deals with these topics it is in order to show that they are philosophical puzzles which are raised from questionable frames, but are ultimately nonsensical [62]. In Hutto and Myin (2013, p.68 ff.), this strategy assumes the form of an out-and-out argument, but the point is the same: to show that, once seen from a different perspective, "hard problems" of this kind dissolve. At best, they serve the function of exposing the flaws of traditional ways of thinking and the need for different accounts [63]. A novel account of phenomenality, coherently, would give up the notion of *qualia* and allow for different dimensions: "Enactivists foreground the ways in which environment-involving activities are required for understanding and con¬ceiving of phenomenality" [64]. If this is true for basic minds, it also holds for complex minds. More complicated environment-involving activities – and not more complicated *qualia* – are enough for understanding and conceiving the phenomenal character of human experiences in general, including the understanding of music, the enjoying of a conversation, the recognition of a reproach in a look, and generally the imponderable evidence which regulates social life.

5   A conclusion with two more clues

I hope to have shown, in what precedes, both why Wittgenstein cannot be considered an *ante-litteram* enactivist, and why – his negative attitude towards science notwithstanding – methodological reflection inspired by his philosophy can be of help for cognitive sciences. Enactivism is a paradigm which fits in with his way of describing psychological concepts and with his outlook on classical problems such as that of mental representation and the inner/outer dichotomy. But other more radical objections to cognitivism, and to some forms of enactivism as well, are suggested by his stance with respect to the "explanatory gap" and the "hard problem" of consciousness.

In concluding this paper, I would like to add two further clues which can be derived from his work.

The first one has to do with the debate regarding the contentful vs. noncontentful and the conceptual vs. nonconceptual character of basic cognition. I suggest that Wittgenstein's work on *aspect-seeing*[65] be taken fully into account in this re-

---

[62]Hutto (2006b, p.14 ff.), Hutto (2006a).
[63]Hutto and Myin (2013, p.169).
[64]Hutto and Myin (2013, p.177).
[65]See also Wittgenstein (2001, sec.XI, part II. p.165).

spect. His remarks are usually considered an important piece of evidence in favor of the conceptual and contentful nature of perception. But what risks being overlooked is his invitation to compare cases, see differences, discriminate between apparently akin situations. This Wittgensteinian explorations, it seems to me, could provide an interesting terrain for a fruitful and not dichotomist debate.

The second clue regards enactivists' recent attempts to extend their domain from low-level to high-level cognition[66]. How can an approach which focuses mainly on basic bodily skills account for more complex, intersubjective, social, cultural phenomena? What is probably needed is a closer examination of what links the two poles of – roughly – body and culture; and what links them is *acting according to rules*, that is, practices. Two Wittgensteinian tools, again, may be of help: his remarks about following a rule, and his conception of the background. But this, of course, is a topic for a wider inquiry.

---

[66]I am referring here to the already rich literature about participatory sense-making (De Jeagher and Di Paolo 2007). An interesting criticism can be found also in Steiner and Stewart (2009), situated normativity (Rietveld 2008), cultural enactivism (Baerveldt and Verheggen 2012), and to Hutto's Narrative Practice Hypothesis (Hutto 2008).

References

Baerveldt, C. and T. Verheggen (2012). Enactivism. In J. Valsiner (Ed.), *The Oxford Handbook of Culture and Psychology*. Oxford Univeristy Press.

Bennett, M. and P. Hacker (2003). *Philosophical Foundations of Neuroscience.* Oxford: Blackwell.

Bitbol, M. (2002). Science as if Situation Mattered. *Phenomenology and the Cognitive Sciences 1*(2), 181–224.

Boncompagni, A. (2012a). Streams and Riverbeds. James' Stream of Thought in Wittgenstein's Manuscripts 165 and 129. *European Journal of Pragmatism and American Philosophy 4*(2), 36–53.

Boncompagni, A. (2012b). *Wittgenstein. Lo sguardo e il limite.* Milano: Mimesis.

Cappuccio, M. and M. Wheeler (2012). Ground-level intelligence: Action-oriented representation and the dynamics of the background. In Z. Radman (Ed.), *Knowing without Thinking*. Macmillan.

Chalmers, D. (1995). Facing up to the Problem of Consciousness. *Journal of Consciousness Studies 2*(3), 200–219.

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

Coates, P. (2007). Experience, Action and Representations: Critical Realism and the Enactive Theory of Vision. *Phenomenology and the Cognitive Sciences 6*(4), 445–462.

Cole, J. (2007). Wittgenstein's Neurophenomenology. *Medical Humanities 33*(1), 59–64.

De Jeagher, H. and E. Di Paolo (2007). Participatory sense-making. An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences 6*(4), 485–507.

Dreyfus, H. (2002). Intelligence without representations. *Phenomenology and the Cognitive Sciences 1*(4), 367–383.

Engel, A. (2010). Directive minds: How dynamics shapes cognition. In J. Stewart, O. Gapenne, and E. Di Paolo (Eds.), *Enaction. Toward a New Paradigm for Cognitive Science*. MIT Press.

Fischer, E. (2011). *Philosophical Delusion and its Therapy. Outline of a Philosophical Revolution.* London: Routledge.

Gallagher, S. (2010). Neurophilosophy and neurophenomenology. In L. Embree and T. Nenon (Eds.), *Phenomenology 2005 (Vol. 5)*. Bucharest: Zeta Press.

Goodman, R. (2002). *Wittgenstein and William James*. Cambridge: Cambridge University Press.

Hutto, D. (2005). Knowing *what*? Radical versus Conservative Enactivism. *Phenomenology and the Cognitive Sciences 4*(4), 389–405.

Hutto, D. (2006a). Turning Hard Problems on their Heads. *Phenomenology and the Cognitive Sciences 5*(1), 75–88.

Hutto, D. (2006b). Unprincipled engagements. In R. Menary (Ed.), *Radical Enactivism. Focus on the Philosophy of Daniel D. Hutto*. Amsterdam: John Benjamins.

Hutto, D. (2008). *Folk Psychological Narratives*. Cambridge: MIT Press.

Hutto, D. and E. Myin (2013). *Radicalizing Enactivism. Basic minds without Content*. Cambridge: MIT Press.

James, W. (1892). *Psychology: Briefer Course*. Cambridge: Harvard University Press.

Levine, J. (1983). Materialism and *qualia*: the Explanatory Gap. *Pacific Philosophical Quarterly 64*, 356–361.

Mandik, P. (2005). Action oriented representations. In A. Brook and K. Akins (Eds.), *Cognition and the Brain: the Philosophy and Neuroscience Movement*. Cambridge: Cambridge University Press.

Menary, R. (2010). Introduction to the Special Issue on 4E Cognition. *Phenomenology and the Cognitive Sciences 9*(4), 459–463.

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

O'Regan, J. and A. Noë (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behavioural and Brain Sciences 24*(5), 939–1031.

Perissinotto, L. (1991). *Logica e immagine del mondo*. Milano: Guerini.

Petitot, J., F. Varela, B. Pachoud, and J.-M. Roy (1999). *Naturalizing Phenomenology*. Stanford: Stanford University Press.

Rietveld, E. (2008). Situated Normativity: the Normative Aspect of Embodied Cognition in Unreflected Action. *Mind 117*(468), 973–1001.

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

Rowlands, M. (2007). Understanding the 'active' in 'enactive'. *Phenomenology and the Cognitive Sciences 6*(4), 427–443.

Ryle, G. (1945). Knowing How and Knowing That: the Presidential Address. *Proceedings of the Aristotelian Society 46*, 1–16.

Schulte, J. (1995). Emotions: Remarks on wittgenstein and james. In R. Egidi (Ed.), *Wittgenstein: Mind and Language*. Klower Academic Publishers.

Stanley, S. and T. Williamson (2001). Knowing How. *Journal of Philosophy 98*(8), 411–444.

Steiner, P. and J. Stewart (2009). From Autonomy to Heteronomy (and Back): The Enaction of Social Life. *Phenomenology and the Cognitive Sciences 8*(4), 527–550.

Stewart, J. (2010). Foundational issues in enaction as a paradigm for cognitive science: From the origins of life to consciousness and writing. In J. Stewart, O. Gapenne, and E. Di Paolo (Eds.), *Enaction. Toward a New Paradigm for Cognitive Science*. Cambridge: MIT Press.

ter Hark, M. (2004). 'patterns of life': a third wittgenstein's concept. In M. Sharrock (Ed.), *The Third Wittgenstein*. Adelshot: Ashgate.

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

Tripodi, P. (2009). Wittgenstein e il naturalismo. *Etica Politica / Ethics Politics 11*(2), 121−141.

Varela, F. (1996). Neurophenomenology. A Methodological Remedy for the 'Hard Problem'. *Journal of Consciousness Studies 3*(4), 330–349.

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

Wittgenstein, L. (1953/2001). *Philosophical Investigations*. Oxford: Blackwell.

Wittgenstein, L. (1958/1969a). *The Blue and Brown Books*. Oxford: Blackwell.

Wittgenstein, L. (1969b). *On Certainty*. Oxford: Blackwell.

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*. Oxford: Blackwell.

Wittgenstein, L. (1982). *Last Writings on the Philosophy of Psychology, Vol. 1.* Oxford: Blackwell.

Wittgenstein, L. (1988). *Wittgenstein's Lectures on Philosophical Psychology*. Brighton: Harvester. edited by P.T. Geach.

Wittgenstein, L. (1992). *Last Writings on the Philosophy of Psychology, Vol. 2.* Oxford: Blackwell.

Wittgenstein, L. (2000). *Wittgenstein's Nachlass. The Bergen Electronic Edition.* Oxford: Oxford University Press.