

THE EXTENDED MIND HYPOTHESIS: AN ANTI-METAPHYSICAL VACCINE

A HIPÓTESE DA MENTE ESTENDIDA: UMA VACINA ANTI-METAFÍSICA

GIORGIO AIROLDI¹

Universidad Nacional de Educación a Distancia (UNED) – Spain
gairoldi1@alumno.uned.es

ABSTRACT: Discussions about the extended mind have ‘extended’ in various directions in the last decades. While applied to other aspects of human cognition and even consciousness, the extended-mind hypothesis has also been criticized, as it questions fundamental ideas such as the image of a dual world, divided between an external and an internal domain by the border of ‘skin and skull’, the idea of a localized and constant decision center, and the role of internal representations. We suggest that the main virtue of the hypothesis is not as a theory per se, but as a vaccine against persistent metaphysical prejudices about the mind’s structure, functions and borders. Being an hypothesis about the most efficient ways to combine resources and problems, and not a theory about the mind’s a-priori constitution, the extended mind view moves the focus from ontology to pragmatics and helps purify philosophy of mind from metaphysical remainders.

KEYWORDS: Extended mind. Externalism. Internalism. Causal-constitutive fallacy.

RESUMO: *As discussões sobre a mente estendida “se estenderam” em várias direções nas últimas décadas. Embora aplicada a outros aspectos da cognição humana e até à consciência, a hipótese da mente estendida também tem sido criticada, pois questiona ideias fundamentais como a imagem de um mundo duplo, dividido entre um domínio externo e um interno pela fronteira da “pele e do crânio”, a ideia de um centro de decisão localizado e constante, e o papel das representações internas. Sugerimos que a principal virtude da hipótese não é uma teoria per se, mas uma vacina contra preconceitos metafísicos persistentes sobre a estrutura, funções e fronteiras da mente. Sendo uma hipótese sobre os modos mais eficientes de combinar recursos e problemas, e não uma teoria sobre a constituição a priori da mente, a visão da mente estendida move o foco da ontologia para a pragmática e ajuda a purificar a filosofia da mente de resquícios metafísicos.*

PALAVRAS-CHAVE: *Mente estendida. Externalismo. Internalismo. Falácia causal-constitutiva.*

INTRODUCTION

Since the publication of “The Extended Mind” (CLARK & CHALMERS, 1998), discussions around the extended mind have flourished and ‘extended’ in various directions. On the one hand, the extended cognition hypothesis, in what is known as its ‘the second wave’, has been developed and fine-tuned, and applications of its principle have been proposed for other aspects of human cognition like

¹ Department of Logic, History and Philosophy of Science at Universidad Nacional de Educación a Distancia (UNED).

perceptions, emotions, and even consciousness. On the other hand, radical criticisms of the concepts upon which the hypothesis is based (e.g. the definition of cognitive, its explanatory power, etc.) have led to qualifying and in some cases to reformulating some of its facets. Nevertheless, almost two decades after it was first proposed, the hypothesis seems to be in wonderful shape. This is to be ascribed to the fact that it addresses some of the deepest issues in philosophy of mind, such as the differences between internal and external, the function of representation, and the ontology of mind; and to the fact that, thanks to the diatribes that it fosters, it represents a powerful vaccine against metaphysical temptations and a constant incentive to the progress of research.

In this paper, we present first of all Clark and Chalmers' original idea, in order to illustrate the main hypothesis upon which it rests. In the second section, we briefly show the radical differences between the focus of the first and the second wave of arguments. The third section presents some of the proposals of extension of the hypothesis to other areas of cognition. In the fourth section, we propose a classification of the different stances about the mind that oppose each other in various degrees (internalism, embodied and embedded perspectives, etc.). Finally, we list the main discussions that arise with respect to the causal-constitutive fallacy, the mark of the cognitive, the difference between internal and external processes, and the explanatory power of the hypothesis of extension. The paper finishes with some reflections and conclusions around what we consider to be the virtues of the hypothesis of the extended mind and the reasons for its success.

1 INTERNALISM, EXTERNALISM AND THE EXTENDED MIND

The extended mind hypothesis is a development of externalism. According to externalism, the contents of intentional states are relational properties identifiable only with reference to some reality (physical, linguistic, social, or metaphysical) external to the individual. On the opposite side, internalism considers intentional states as intrinsic properties of the individual, independent from the external reality (BUZZONI, 2006).² The 'classical' externalism initially proposed by Putnam (PUTNAM, 1976) and Burge (BURGE, 1979) does not, nevertheless, grant any active role to external elements in the determination of the intentional content: these play a purely passive role in cognitive processes and in the definition of the mental states of the individual.

Clark and Chalmers' extended mind hypothesis introduces a new type of externalism, in which external elements do have an 'active' causal role (CLARK & CHALMERS, 1998).³ The central issue in the Otto case (CLARK & CHALMERS, 1998)⁴ is not to counter a 'broad' versus a 'narrow' content, as it would be the case with

² The debate between internalism and externalism was initially fostered by the thought experiment of Twin Earth, proposed by Putnam (1975). Burge (1979) extends the externalist hypothesis and claims that mental content depends upon the social environment. For an overview of linguistic externalism see Bezuidenhout (2008).

³ "[...] the relevant external features are *active*" (CLARK & CHALMERS, 1998, p. 09).

⁴ In the article, the case of Otto is described, a patient with Alzheimers who stores his beliefs in a portable notebook that he always keeps with himself.

classical externalism, but to defend a narrow content extended to the environment (CLARK, 2010a).⁵ An example of cognitive phenomenon of this kind is the act of gesturing while talking, an activity which is corporal and neuronal at the same time: verbal thoughts and physical gestures influence each other and constitute a coupled system. Goldin-Meadow (2003) suggests that gesture can provide an alternative representational format that adds information in either an analog, motoric, or visuospatial way. Additionally, gesture can reduce the overall neural cognitive load and free resources for other tasks.⁶ Gestures, thus, do not just express internal thoughts complete by themselves, but are part of these thoughts and causally interact with the cognitive system (CLARK, 2011, p. 123-126).

The heart of the issue here is not the trivial observation that external elements causally influence the process (for example, by increasing our memory or by helping us when making calculations). Clark and Chalmers affirm that the causal relationship between external and internal elements, far from being distant and historical, is mutual and diachronic. Clark applies the notion of ‘continuous reciprocal causation’ to cognitive activities that involve “continuous, mutually modulatory influences linking brain, body, and world” (CLARK, 1997, p. 163): playing in an orchestra or having a group conversation are examples of such activities whose explanation cannot be given in terms of input to and output from a closed cognitive system (CLARK, 1997, p. 165). The cognitive system consists of the sum of both kinds of elements (external and internal), and constitutes a real coupled dynamical system (MENARY, 2010a, p. 03-04).

Not all coupled dynamical systems are cognitive, though. Clark and Chalmers propose a fundamental criterion to limit the systems that can be really considered ‘extended minds’: the Parity Principle. The principle affirms that an element⁷ belongs to a cognitive system not because of its localization (for example, within or outside of the brain), but because of its function. The authors list portability, availability, typical and uncritical use, and easiness of access as necessary and sufficient characteristics for an element to belong to the cognitive system: according to these criteria, a notebook would not count as such, while Otto’s portable notebook would (CLARK, 2010a, p. 44-47). The Parity Principle clearly shows the commitment of the original extended mind hypothesis with functionalism⁸, as it identifies and classifies mental states in terms of their causal roles (BECHTEL, 1988).⁹ Together with functionalism’s flaws,¹⁰ the extended mind hypothesis shares its main virtue: it dissociates the cognitive from the physical; a silicon circuit or a Martian organ can be as cognitive as a human neuronal circuit.

⁵ “[...] what was at issue was more like an environmentally extended case of narrow content than a case of broad content” (CLARK, 2010a, p. 45).

⁶ Goldin-Meadow (2003) presents the case of reduced capacity of memorizing a list of words by a group of children not allowed to gesture with respect to another group which could freely gesture.

⁷ The reference to ‘an element’ is important, given that the principle is a criterion of ‘belonging to’ and not of constitution: distinction that, as it will be shown, lies at the basis of the discussion with Adams and Aizawa.

⁸ The ‘second wave’ of the extended mind hypothesis overcomes this dependency.

⁹ “Functionalism maintains that mental events are classified in terms of their causal role” (BECHTEL, 1988, p. 112).

¹⁰ Bechtel (1988, p. 123-136) lists a summary of the main criticisms to functionalism.

There are different views with regards to the reach of this functionalism: Clark considers that the Parity Principle involves a ‘very weak’ functionalism that, according to Chalmers (CLARK, 2011),¹¹ does not extend to consciousness, and that just denies any relevance of the internal/external difference for the cognitive processes; Wheeler defends, on the contrary, an extended functionalism (WHEELER, 2010).¹²

The classical vision considers the mind as a Turing machine with a certain ability for computation and manipulation of symbols, result of the accumulation of adaptive responses. The extended mind hypothesis suggests an idea of human cognition radically different, as a product of the hybridization between the brain and technological artifacts, implying that human beings are ‘natural cyborgs’. There is a continuous range of intermediate cases between behaviors and decisions based on discursive rationality and ‘quasi-automatic’ ones, in which the active contribution of the body, of the social environment, and of cultural artifacts can sometimes prevail over the role of the brain in solving problems, freeing the latter from a relevant working load. The extended mind hypothesis is about reinventing cognition as a distributed capacity (CLARK, 2001, p. 121-129).

Under this view, human cognition springs from the collaboration among body, brain, and the active contribution of the technological environment: the artist draws a sketch before painting the landscape because the sketch is part of the creative process as much as her hand or brain, and not just a simple temporary information store (CLARK, 2001, p. 133).

2 THE TWO ‘WAVES’ OF THE HYPOTHESIS: PARITY AND COMPLEMENTARITY

The internalist view considers that the brain is the only place where cognitive activities take place, and attributes to this organ the capacity of unplugging from the environment and of managing representations. The extended mind hypothesis rejects this possibility by highlighting two mechanisms at the basis of the extension: functionality and complementarity.

The Parity Principle suggests that it is the function, and not the location, what makes something cognitive: an external element belongs to the cognitive system because it plays a functional role identical to the one that an internal element would play. The principle imposes a ‘veil of ignorance’ that makes irrelevant the border of ‘skull and skin’ and the difference between perception and introspection, and thus avoids the ‘bio-chauvinist’ prejudice: portability and availability of a resource are the only fundamental virtues in cognition (CLARK, 2011, p. 78). By assuming an isomorphism between internal and external processes, however, the Parity Principle is unable to capture the differences between *exograms* (external symbols) and *engrams* (cerebral memory):¹³ a

¹¹ “I think that functionalism about consciousness is implausible” (CHALMERS, 2011, p. xv).

¹² “[...] *the parity principle* forges a strong connection between functionalism and ExC [extended cognition hypothesis]” (WHEELER, 2010, p. 248).

¹³ Terminology suggested by Donald (1991, p. 314), who names *exograms* after Lashley’s (1950) *engrams*, single entries in the biological memory system.

checklist and an engram are not equivalent in many ways; moreover, it does not distinguish individual differences with regards to the use of available resources (for example, some people might prefer memorizing information, while others checking it on an agenda).

The Complementary Principle (SUTTON, 2010, p. 194), that opens up the so-called ‘second wave’ of the extended mind hypothesis, overcomes these problems by including an external element in the cognitive system precisely because it plays a different role than an internal one, and that the latter could not perform. By focusing on the cognitive contribution of the specific features of non-neuronal elements, the principle acknowledges that external traits are not isomorphic with internal ones, but complementary: their substantial difference is precisely what gives value to their contribution (SUTTON, 2010). Heersmink (2015) suggests to identify the kind and level of integration between agents and artifacts on the basis of several dimensions, among which: the direction of the flow of information between the elements of the cognitive system (e.g. road signs are one-way from artifact to agents, post-it notes are two-way from agent to artifact to agent); the reliability of the access to external information (e.g. analogue notebooks are more reliable than electronic ones because they do not require electricity); the durability of the relationship with the artifact (e.g. shopping list are on-offs, abacuses are re-used); the trust granted to the information (e.g. to a Encyclopedia Britanica article vs. a Wikipedia entry); the procedural and informational transparency of the cognitive artifact (i.e. the effortlessness in using and interpreting the artifact); the individualization or interchangeability of the artifact (e.g. a underlined book are not interchangeable with other copies); the transformation of the representational and cognitive capacities of the agent by the use of the artifact (as in enculturation, e.g. the brain of a baby modified when learning to speak).¹⁴ Heersmink do not consider that these dimensions constitute necessary and sufficient conditions: they rather “provide a toolbox for investigating the degree and nature of the integration of agent and artifact into ‘new systemic wholes’” (HEERSMINK, 2015, p. 596).

The two approaches (functionalism and complementarism) are not necessarily opposed in all their aspects (KIVERSTEIN & FARINA, 2011), as, some might argue, functional isomorphism is not explicitly required by the principle, as Clark and Chalmers only claim that “[i]f, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process. Cognitive processes ain’t (all) in the head!” (CLARK & CHALMERS, 1998, p. 08).

¹⁴ Heersmink (2015, p. 591) makes the interesting comment that “this is the reversed version of the parity principle”, as “external states and processes, rather than being functionally isomorphic to internal ones, are soaked up by the brain which then takes on the properties of those external representational systems”.

3 *EXTENDING* THE EXTENDED MIND: FROM BELIEFS TO CONSCIOUSNESS AND TO NICHE CONSTRUCTION

The extended mind hypothesis considers that external elements not only play a causal role in intelligent behavior, but also show a constitutive interdependence with it and, in some cases, allow behavior that would not be possible if the cognitive system lacked them (WHEELER, 2010, p. 246-247): mental processes and beliefs can constitutively include environmental resources. The natural extension of the hypothesis points to the possibility that also other intentional mental states, reasoning, perceptions and emotions could extend to the environment.

Clark claims that such an extension is real: for example, when one uses the camera of a mobile phone as extension of her visual capacity, provided the camera fulfills the criteria of the Parity Principle. Chalmers, consistent with his rejection of functionalism in the realm of consciousness, excludes from it also the extended mind hypothesis, given that he does not conceive a thought experiment in which two identical twins can have different self-experiences: he imputes such impossibility to the band-width required by consciousness to access information, band-width not supported by our current perceptual system (CLARK, 2011, p. xiv-xv). Rowlands, on the contrary, claims that consciousness too can be extended to the environment, moving the focus from the definition of consciousness as intentional object of an apprehension act to consciousness as determinant of a reference (ROWLANDS, 2010).

By denying the possibility of an extended consciousness, and at the same time admitting the possibility of extended states of mind, one grants to the brain a special role, and at the same time rejects the identification of the mental with the conscious.¹⁵ Clark solves this apparent dilemma by diluting his externalism with some dose of internalism, given that he keeps within the brain, besides consciousness, many of the functions that internalism traditionally grants it, such as the ability to create representations and the possibility to develop cognitive processes isolated from the environment (that is, off-line, or non-extended): in short, he accepts the internalist view of the brain as the center of the cognitive. However, conversely to classical internalism, Clark claims that the brain, even if it is the center of the cognitive system, does not bother for the localization (either internal or external) of its resources and processes, as long as they are useful to solve problems. The human cognitive system is ‘organism centered’, but not ‘organism bound’. Clark also clearly rejects the internalist view of the brain as ‘final decider’: internalism does not specify which part of the brain, and based on which functions, makes decisions, and it is even doubtful that decisions as such appear in a single point of the cognitive system (CLARK, 2011, p. 159-160).

¹⁵ Clark underscores this repeatedly: “Finally, we allowed that conscious mental states [...] supervene only on local processes inside the head. But insofar as the scope of the mental is held to outrun that of conscious...” (2010a, p. 45; 2011, p. 79); “Shrinking the mind to the conscious is certainly one way to avoid the conclusions of the original paper. But do we really want to shrink the mind so far?” (2011, p. 161).

A more radical interpretation considers the extended mind hypothesis as a particular and extreme case of a general and complex hybridization between living beings and environment (physical, social, technological and cultural). The human being is *wholly extended* into the environment, as it is the case with digestion, linked to artifacts and techniques for the preparation and cooking of food. In the *niche construction model* (STERENLY, 2010), agents modify their environment and are modified by it: many animals build dens, and some trees modify the land where they grow (for example, eucalypts increase the acidity of the ground close to their trunk in order to avoid other plants to take root close to them). Niches are built around certain resources and identified with regards to three axes: the availability and reliability of the resource (external resources, for example, being shared, could be more reliable than internal ones); its standardization (external personalized resources, such as Otto's portable notebook, or a blind person's stick, are placed in one end of this axis, given that they are perceived as part of the body; a book full of hand-notes is placed in an intermediate position, as it is personalized but not incorporated); its shared and collective nature (the mnemonic resources of Elizabethan actors are an example of external shared memory (SUTTON, 2010, p. 208-209)). The extended mind hypothesis is, under this view, just a particular and extreme case of niche construction: the case of highly reliable, incorporated and personalized external resources.

4 DEGREES OF EXTENSION

Between pure internalism and active externalism, there are several postures that differ with regards to the extension of the mental and its relationship with the environment: we briefly list the most interesting features of the main ones.

Pure internalism exists in two fundamental versions: one, defended for example by Adams and Aizawa (2001)¹⁶ and defined as *brainbound* by Clark (2011, p. xxvii), limits cognition to the neuronal; the other, known in the literature as *embodied*, recognizes in its radical version a cognitive and explanatory role also for the body (thus approaching externalism) and, in more moderate versions, to representations, especially to mental representations of the body (ADAMS & AIZAWA, 2001).

A 'weak' externalist position, known as the *embedded hypothesis* (RUPERT, 2004) admits that part of the environment has a causal role in some cognitive processes, although it is not a constitutive part of them (Wheeler 2010: 246). The embedded hypothesis highlights the difference between external and internal states (for example, between biological memory and external memory, CLARK, 2011, p. 112-113) to escape the causal-constitutive fallacy¹⁷ (ADAMS & AIZAWA, 2010, p. 67).

¹⁶ "We maintain that [...] processes that are plausibly construed to be cognitive occur within the brain, do not occur outside of the brain, and do not cross the bonds of the brain" (ADAMS & AIZAWA, 2010, p. 74).

¹⁷ See section 5 for an analysis of the causal-constitutive fallacy.

Finally, the ‘strong’ externalist position, linked to the extended mind hypothesis, considers that environment to be a constitutive part of the cognitive system, that in turn includes the brain, the body and the physical, cultural and social environment, all interconnected through systemic relationships of feed-back, feed-forward, etc. (CLARK, 2011, p. xxviii).

The extended mind hypothesis can be taken to the extreme of denying that the brain generates representations. According to this view, the cognitive system resorts to the environment as a model of itself, without any need of creating internal images. This view derives from Brooks’ suggestion to move from approaches in Artificial Intelligence highlighting abstract manipulation of symbols to a methodology emphasizing ongoing physical interaction with the environment, provided that “the world is its own best model” (BROOKS, 1999, p. 115): if representations are grounded on the physical world, the need for traditional symbolic representations fades entirely. The known incapacity of the subject, if focused on some detail of a scene, to detect even very relevant changes in other parts of it, is considered as an evidence of this claim: if an internal image existed, any change in the world would be detected, once reality were matched with the mental model. As Merleau-Ponty (2012) points out, the body is always present and is thus the best model of itself. The reason for this rejection of representations is that they are seen as a return to mentalism, as they weaken the causal role and explanatory force of the external element (ALSMITH & DE VIGNEMONT, 2012).¹⁸ There are hypotheses, though, that avoid this risk without coming to such an extreme conclusion (SIMONS & RENSINK, 2005): the representation could, for example, arise and decay quickly, or persist without being taken into account. As a result, the phenomenon of ‘blindness to details’ might imply not that there are no representations of the environment, but rather that there are several, partial ones, not so easily available to conscious access. Moreover, for the principle of minimum effort, the subject would rather access directly the environment, when available, to gain information, even if an internal representation of it were available (CLARK, 2011, p. 141-146).

The following table summarizes the degree of ‘external extension’ of the cognitive system defended by the discussed postures: internalism and the extended mind hypothesis just represent the two extreme positions of a rich range of options.

¹⁸ “[...] whether positing body representations actually undermines the explanatory role of the body” (ALSMITH & DE VIGNEMONT, 2012, p. 02).

Brain	Body		Environment	
	Representation	Real	Causal role	Inter-connected role
Internalism/ Brainbound				
Moderate Embodied				
Radical Embodied				
Embedded or Scaffolded				
Extended Mind				

Table 1 – Degrees of extension of the cognitive system

5 FRUITFUL CONTROVERSIES

The great amount of discussions and fruitful controversies that the extended mind hypothesis has not ceased to foster since its appearance in 1998 is perhaps the best evidence of its importance for philosophy of mind and of the topicality of the issues that it addresses. The main ones are presented in this section.

The causal-constitutive fallacy

Adams and Aizawa defend an internalist position that locates all cognitive processes within the brain, and reject the extended mind hypothesis by denying that it is sufficient for an external element to play a causal role in such processes to be part of them: believing it means falling in the causal-constitutive fallacy (ADAMS & AIZAWA, 2010, p. 73-76).

Answers to this criticism are numerous and assorted: Clark and Menary discard it; Palermos recognizes that it does have some ground and suggests, to escape it, a modification of the extended mind hypothesis; Ross and Ladyman claim that it makes no sense at all. Let's briefly go through each of these positions.

Clark claims that the issue is not to know whether an element is *in itself* cognitive when used by a cognitive system (or one would fall in the 'compositional fallacy', the assumption that the parts of a system must have the same characteristics than the system itself). The question is rather whether an element is part of a cognitive system (CLARK, 2010ab, p. 82-85). Adams and Aizawa's position presupposes that there is a well-defined cognitive agent to which an external

element couples just because it has a casual impact on it: but such an agent does not exist, and the cognitive system is constituted by the whole set of all elements intervening in it. The definition of 'extended' itself might perhaps suggest the idea of a cognitive center: 'distributed' mind would be a more appropriate definition (MENARY, 2010b, p. 606-610).

Palermos accepts that the original version of the extended mind hypothesis exhibits the causal-constitutive fallacy (PALERMOS, 2014).¹⁹ He proposes to apply the theory of dynamical systems²⁰ to the modelling of cognitive processes in order to avoid the identification of causal and constitutive, and to supplant it with the stronger and more defensible identification of dynamic and non-linear inter-causality. In this way, he claims that the extension hypothesis is saved and the fallacy avoided.

Ross and Ladyman reject discussion about the fallacy because they consider it to be based on metaphysical principles. Questions about the inclusion of an element in a system because of its causal role are senseless, given that neither the concept of system, nor the concepts of causality or constitution have a place in a mature science (ROSS & LADYMAN, 2010).

The mark of the cognitive

The abundance of intermediate postures between internalism and the extended mind hypothesis (Table 1) is in part due to a lack of a universal definition of cognition. Each view develops around a more or less implicit interpretation of what mind is: it is therefore necessary to make it explicit, to avoid using a proposal about the limits of the cognitive system to deduce a definition of cognition that lays at the basis of the initial proposal, plunging into the pitfall of a *petitio principii*. If cognition includes activities such as remembering, perceiving, learning, or reasoning, what do they all share that makes them cognitive? The range of answers is very wide, from the classical computationalism that identifies cognitive vehicles with symbols equipped with syntactic properties; to connectionism that identifies them with algorithms of nodes' activations in a network. In a different line, Wilson (2010, p. 183) moves the focus from representation to the act of representing as a mark of the cognitive, and Rowlands identifies a cognitive process in whatever allows the performing of a cognitive action. Alternatively, the cognitive integration approach acknowledges that our primary cognitive engagements with the world are embodied and primarily sensorimotor ones. It thus integrates the bodily 'internal' and 'external' facets of cognition, and understands this integration in terms of manipulation of environmental vehicles (MENARY, 2010a, p. 268-269). The integrationist view includes the linguistic and representational environment within the cognitive system, given that they allow cognitive actions impossible based on neural activity alone (MENARY, 2010b, p. 611). Sometimes, the distinction between the mental and the cognitive, and how the extended mind hypothesis is reflected

¹⁹ "Such a 'cognitive bloat' would actually be the outcome of repeatedly committing the 'causal-constitution' fallacy that Adams and Aizawa pointed out" (PALERMOS, 2014, p. 07).

²⁰ See next section for a brief description of the theory.

in each, are also poorly defined. Carter and colleagues (CARTER et al., 2014), for example, do explicitly distinguish among extended *cognition*, or the claim that *cognitive processing* can be suprapersonal, extended *mind*, or the claim that *mental processes* can be suprapersonal, and distributed cognition, or the claim that cognitive processing can be distributed across several agents and artifacts. In view of such variety, let's review the main postures regarding the mark of the cognitive.

Adams and Aizawa reduce the cognitive to the manipulation of representations with 'non-derived' content, located uniquely in the brain. What is external to the brain, processes and representations, has always a derived content (for example, the meaning of alphabetical signs derives from social norms): evidence of this is that we can identify an external representation and socially agree to change its content (for example, that the meaning of green for traffic lights becomes 'go'), but we cannot do the same with internal representations (because we do not know which neuronal networks identify them,²¹ and we would anyhow not have the ability to manipulate them)²². Aizawa and Adams suggest that "there is nothing in DNA or its causal activity during development that resembles the way that meaning is assigned by a human mind to an artifact or Symbol" and that "the derivation of the human mind from the human genome is unlike the derivation of derived content from prior content" (2005, p. 667): it is, therefore, non-derivative. There are many criticisms of this position. In the first place, the authors neither supply an exact definition of 'representation with non-derived content', nor a definition of what differentiates it from one with intrinsic content (CLARK, 2010a, p. 90). Moreover, the proposal seems to be a definition rather than a fact, given that the authors do not justify the view that the cognitive is characterized by intrinsic content: they even claim that it is not clear to what extent cognitive states could also involve derived contents.²³ Dennett denies the existence of original (or non-derived) intentionality by asking "[w]here, though, do we get our 'original' and underived intentionality? From God, as Michelangelo suggests?" (DENNETT, 1990, p. 54).

According to Clark, a process is cognitive if it supports intelligent behavior. The focus is substantially different than in Adams and Aizawa: the mark of the cognitive does not lay in the nature of an element (derived or non-derived content), but depends on its function within the system. Otto's notebook does not (and can not) have any non-derived content, but it is functionally linked to Otto's cognitive system in a dispositional way, in a coupling defined by the Parity Principle. The cognitive is not an intrinsic feature of any isolated element, it rather supervenes on a system whose elements as a set form a cognitive process.

Hurley claims that the putative 'causal-constitutive fallacy' simply masks the prejudice of identifying cognitive with 'internal', while limiting the external role to a causal contribution (HURLEY, 2010, p. 126). This posture considers the internal

²¹ "We don't know what specific syntactic item in the brain bears that content" (ADAMS & AIZAWA, 2010, p. 73).

²² "[...] we have no way to identify particular tokens of brain states qua syntactic items in order to affix contents to them" (ADAMS & AIZAWA, 2010, p. 72).

²³ "It is unclear to what extent each cognitive state of each cognitive process must involve non-derived content" (ADAMS & AIZAWA, 2001, p. 50).

and the external as separate domains with fixed properties: but technology, for example, is not necessarily something external, given that learning a sentence by heart means creating a ‘mental artefact’ (SUTTON, 2010, p. 207-208). Moreover, there might exist sets of neurons with just a causal role, even if they constitutively belong to the brain.

Butler (1998, p. 205) identifies the cognitive with the biological brain, because it is where the computational control happens: it rejects the idea that external processes could be cognitive. There are several answers to this view: Clark criticizes it first of all because it does not clarify where exactly decisions would be made; on the other hand, there are zones of the brain that do not take part in the decision process, thus according to this definition they would not be cognitive either; finally, the functional role of the biological memory and of the external memory is identical, so it is unclear why the former would be cognitive and the later would not. Even if a place where the final decision is made existed, it is not clear why it should be the brain (CLARK, 2010a, p. 55-56).

Grush (2003) identifies the mark of the cognitive with the capacity of the brain to control motor activities, and not with its capacity to generate representations, that are just ‘working tools’. Control of some motor activities is possible thanks to feed-back from the environment; given that feed-back’s perception might experience a delay, the cognitive system produces inner dynamical models of the environment that allow to simulate its answers. Representations are nothing more than these off-line models that are used as surrogates of the real environment. Cognitive systems can thus work at the same time by interacting with the real world and with models of it: mind is dis-engaged from the world, but not dis-embodied from the body.

Palermos (2014) criticizes Clark’s Parity Principle as the criterion to define the cognitive. Interpreting it as ‘glue and trust’ principle,²⁴ he claims that it allows the inclusion of almost any external element, causing a ‘cognitive bloat’ and falling in the causal-constitutive fallacy. As an alternative, he suggests an interpretation of the cognitive based on the theory of dynamic systems (DST). Contrary to the sequential computation of the classic theory, in which the temporal factor is irrelevant, DST models the cognitive system as the sum of coupled systems, whose mutual and continuous interactions can be described only through non-linear parameters. The advantage of this proposal is that it entails a concept of the cognitive as supervenient and not linked to any isolated element. Clark seems anyhow not to limit his criterion to the Parity Principle and to suggest, in the same

²⁴ The ‘glue and trust’ principle consists in a set of conditions under which cognitive processes should be implemented outside the body in order to represent extended cognition: in particular, that “the resource be reliably available and typically invoked [...] [t]hat any information thus retrieved be more-or-less automatically endorsed [...] [t]hat the information contained in the resource be easily accessible as and when required” (CLARK, 2010a, p. 46). These are just a set of sufficient but not necessary conditions for extended cognition, as other sets can be proposed (AIZAWA, 2015). The equivalence between the Parity Principle and the ‘glue and trust’ principle can thus be questioned, but Palermo’s criticism is not invalidated by this clarification.

line of Palermo's proposal, the need of a strong interconnection among the elements that constitute the cognitive system (CLARK, 2011).²⁵

Finally, Ross and Ladyman do not suggest any criterion of the cognitive, but consider that any aspiration of doing so is metaphysical (ROSS & LADYMAN, 2010). The concept of 'constitutive' belongs to a vision of the world that implies the existence of fundamental elements, out of which all the remaining is composed: it is a vision overcome by mature sciences. The question about the border of the cognitive is to be rejected independently of the answer, because it is an ontological question: even the externalist proposal based on the extended mind hypothesis fosters a vision of the world full of spatially identifiable objects.

In conclusion, the various proposals around the mark of the cognitive as well as the controversies related to the causal-constitutive fallacy, and above all Ross and Ladyman's rejection of any criterion as metaphysical, suggest that the concept of mind itself is superfluous in the cognitive sciences.²⁶ The predicate 'mental' is applied to such a wide variety of realities (from thermostats²⁷ to complex systems within human beings), that the suspicion arises we are not able to recognize a mind when we see one (CLARK, 2010a, p. 62-64).

Internal and external processes

Several criticisms to the extended mind hypothesis highlight the differences between internal and external processes.

Dartnall stresses that the internal biological memory, contrary to the external, is not just a simple informational store: remembering is mainly *creating* information. Clark answers that the memory of Otto's notebook is not cognitive in itself, in the same way as a group of neurons is not a group that can likewise store information in a passive way (for example, when we learn by heart a sentence in a language we do not understand) (CLARK, 2010a, p. 52-53).

Butler (1998, p. 211) emphasizes that Otto's access to information requires some perception, while Inga's access to her memory is fully internal.²⁸ Clark answers that Otto and his notebook constitute a cognitive system *as a set*, therefore the whole process remains within it. Davies (apud CLARK, 2010a, p. 57) adds that Otto could make mistakes while reading, but Clark emphasizes that Inga too could remember wrongly. Finally, the objection that Otto's notebook is public and Inga's memory is private is rejected by Clark by appealing to cases of multiple personalities sharing the same memories (CLARK, 2010a, p. 57-58).

²⁵ "Coupling alone is not enough [...] these are the cases when we confront a recognizably cognitive process, running in some agent, that creates outputs (speech, gesture, expressive movements, written words) that, recycled as inputs, drive the cognitive process along" (CLARK, 2011, p. 131).

²⁶ "[...] a fully general theory of cognition [...] need incorporate no single overarching account of limits on the boundaries of cognitive systems" (ROSS & LADYMAN, 2010, p. 156).

²⁷ See e.g. Dennett (1987).

²⁸ Inga is cited as an example of a person with a memory not affected by Alzheimer (CLARK & CHALMERS, 1998).

Sterelny (2004, p. 246) highlights a substantial difference between perception of the external and introspection: the former can be manipulated. The susceptibility to external attacks differentiates the two kinds of channels. Clark, even if he admits that doubts about the reliability of the information stored in his notebook would cause its decoupling from Otto's cognitive system, claims that the same problem can happen in the biological memory, in which a psychologist could generate false beliefs (CLARK, 2010a, p. 60-61).

Critics of the extended mind hypothesis have also insisted that Otto's access to the information stored in his notebook entails two steps (he believes that his belief is in the notebook and then he retrieves it), while Inga's access to her memory is direct. Moreover, Inga can claim to have a 'first person' authority upon her internal beliefs, while Otto has to find out what it is that he believes, by checking in his notebook (PRESTON, 2010). Clark answers that the use of his notebook has become so automatic for Otto that he does not even notice it anymore, the same as Inga's memory: it's a tool in Heidegger's sense (HEDEGGER, 1996; SCHMITT, 1965).

Finally, Rupert claims that a 'classical' cognitive system is persistent, while an extended one is brief: it only works as long as its external elements are available (RUPERT, 2010). Clark answers that the spider's web is not always available either, but this does not hinder the spider's hunting system to be constituted by both spider and net.

Explanatory power

Another criticism of the extended mind hypothesis points out that it does not provide greater explanatory power than the internalist view; moreover, internalism has allowed much progress in the cognitive sciences that should not be abandoned without a clearly better alternative. Adams and Aizawa advocate for keeping a clear separation between the internal field, where interesting regularities have been identified, and the external one, where the enormous variety of phenomena seems to reduce the possibility of finding general laws (ADAMS & AIZAWA, 2010).

Clark rejects this distinction because he believes that the differences between internal and external processes are not greater than those among internal processes (CLARK, 2010a, p. 51). Moreover, Sutton emphasizes that rejecting the study of external cognitive processes because of their variety is like rejecting the study of the nature of mirrors by focusing on the variety of the images they reflect (SUTTON, 2010, p. 214).

Finally, the request for unification made by Adams and Aizawa is fulfilled, given that the extended mind hypothesis really *unifies* explanations about the working of cognitive systems (CHALMERS, 2011, p. xiv).

CONCLUSIONS

The multiplicity of the issues under discussion and of the controversies surrounding the extended mind hypothesis are good witnesses of its fecundity. Rather than in the enunciation of a theoretical hypothesis, however interesting and revolutionary, we believe that the main virtue of the idea that some of the cognitive processes can take place outside of the brain consists in its making explicit and questioning prejudices concerning what the mind is, what functions it performs and how, and where its center and its border are located. Many of these issues revolve around the metaphysical view of mind as a substance. From this perspective, the extended mind hypothesis addresses many of the most fundamental areas of philosophy of mind and acts as a powerful vaccine against the danger of more or less explicit metaphysical infections.

First of all, it shows the persistent difficulty in getting rid of the Cartesian dual world, divided between an external and an internal domain, each characterized by its own and incommensurable laws, where, within the border of 'skin and skull', there resides the 'interior homunculus', who contemplates and manipulates the representations that appear in the 'mirror of nature'²⁹ that is the mind. The metaphysics of Cartesian dualism, if endorsed, prevents the understanding of the problem of mental cognition and of mental structure (KENNY, 1992). Radical internalist views, such as Adams and Aizawa's, with their doubtful definition of the cognitive as 'representations with non-derived content', still show a close proximity to a quasi-magical idea of the internal: because they expel from the realm of the mental all acts not based on representations and decouple the mind from its environment, that cannot generate them; and because they claim that only internal processes can be classified and formalized in laws, contrary to the chaotic external world, putatively full of dissimilar phenomena. The internalist proposal can be successfully applied in some cases, but it should be considered an hypothesis useful as a theoretical ideal and not as the description of an ontological reality.

The questioning of the centrality of the internal has important consequences in two other aspects of cognition: the existence of a defined and constant center where decisions are made, and the role of representations. Let's see how.

Regarding the belief that there exists a central mind that makes all the decisions (or, at least, the important ones), already questioned by Dennett (1991), the idea of a cognition distributed and extended to the environment compels us to reflect about why the distributed structure should stop at the limits of the brain and be centralized. Even if one claims, as discussed in the previous paragraph, that all the cognitive states and processes reside in the brain, one can still admit that decision-making is distributed. On the other hand, with regards to the putatively fundamental role of representations in cognition, we have seen that the active role of external resources suggests the possibility of automatic rational actions not based on representations, and of cognitive processes that give priority to the direct access to the world, when it is available. The role of internal representations in problem

²⁹ According to the definition in Rorty (1979), where he defends an anti-representationalist posture.

solving is also doubtful, as long as cognition is active and distributed: representations would play a role similar to the one of external elements (CLARK, 2001, p. 129-131).

Besides, the extended mind hypothesis moves the focus from ontology to pragmatic considerations, as it is an hypothesis about the most efficient ways to combine resources and problems, and not a theory about the *a-priori* constitution of the mind. The Parity Principle is not a definition of the cognitive (as a matter of fact, Clark doesn't engage with any specific definition) and it just identifies a criterion of belonging: instead of establishing first what the cognitive is in a static and ontological way (as if the cognitive were a natural kind)³⁰, and afterwards deciding whether an element that takes part in a cognitive process fulfills the criteria, Clark defines first of all under which conditions an element can be considered part of a cognitive system, without worrying about its ontological status. This posture seems reasonable, because nothing prevents some neural networks from playing a cognitive role in some cases. If, like in the case of 'blindness to details', the brain, even when an internal representation is available, directly resorts to the environment should this make its task easier, there is no reason why it should worry about the ontological status of a resource, as long as the resource can perform a given function. Thanks to its impartiality, the brain seems to believe less than some authors about the 'exceptionality' of the interior.

Table 2 summarizes and contrasts the central points of the discussions between the internalist views and the extended mind hypothesis.

Internalism	Extended Mind
Internal/external dualism	Continuity between external and internal
Clear border of the cognitive (skull & skin, body...)	Border of the cognitive variable depending on the process
Representationalism	Opportunism: representation or reality depending on convenience
Internal laws can be formalized, external laws cannot	Same complexity/variety of internal and external phenomena
Controlling center localized and stable	There is no controlling center, decisions distributed according to variable criteria
Ontology first	Function first

Table 2- Main features of internalism versus extended mind hypothesis

³⁰ In the sense of Quine (1969).

Discussions fostered by the extended mind hypothesis thus represent an exceptional vaccine that helps purify philosophy of mind from metaphysical infections. Defining a concept of cognitive or of mind can be helpful, either as an assumption to limit the scope of an investigation, or as a guideline to organizing theories. On the other hand, searching for an absolute definition does not seem to be more justified than the metaphysical love of taxonomies. Under this instrumental view, none of the two analyzed positions is absolutely truer or better than the other: each can represent a useful tool. One should therefore look not for a clash but rather for a fusion between them, an 'ecumenical' attitude that avoids slipping towards suspicious extremes: if, on one hand, the hypothesis of embedded cognition runs the risk of seeing the external as purely instrumental to the internal, on the other hand, the extended mind hypothesis might forget that the center of human cognition is the organism.

The sum of these views helps us to remember that the target of all research is not the creation of theories but rather the understanding of reality, and that reality is always more complex and elusive than any ontological coat or metaphysical hat we may have for it.³¹

REFERENCES

ADAMS, Fred; AIZAWA, Ken. The bounds of cognition. *Philosophical Psychology*, v. 14, n. 1, pp. 43-64, 2001.

_____. Defending the bounds of cognition. In: MENARY, Richard A. (ed.). *The Extended Mind*. Cambridge: MIT Press, 2010. pp. 67-80.

AIZAWA, Ken. *Extended cognition. Trust and glue, and knowledge*. Text of the talk at the First Annual Extended Knowledge at the University of Edinburgh, April 22-23, 2015. Available at: https://www.academia.edu/12094812/Extended_Cognition_Trust_and_Glue_and_Knowledge. Accessed in: 2019-03-02.

AIZAWA, Ken; ADAMS, Fred. Defending non-derived content. *Philosophical Psychology*, v. 18, n. 6, pp. 661-669, 2005.

ALSMITH; Adrian J. T.; DE VIGNEMONT, Frédérique. Embodying the mind and representing the body. *Review of Philosophy and Psychology*, v. 3, n. 1, pp. 1-13, 2012.

BECHTEL, William. *Philosophy of mind: an overview for cognitive science*. New York, London: Psychology Press, 1988.

BEZUIDENHOUT, Anne L. Language as Internal. In: LEPORE, Ernest; SMITH, Barry C. (eds.). *The Oxford Handbook of Philosophy of Language*. Oxford: Oxford University Press, 2008. pp. 127-139.

³¹ I would like to thank the two anonymous referees whose comments and suggestions have contributed to substantially improving upon the earlier version of this paper.

- BROOKS, Rodney A. *Cambrian intelligence. The early history of the new AI*. The MIT Press, 1999.
- BURGE, Tyler. Individualism and the mental. *Midwest Studies in Philosophy*, v. 4, n.1, pp. 73-122, 1979.
- BUTLER, Keith. L. *Internal affairs: a critique of externalism in the Philosophy of Mind*. Dordrecht: Kluwer, 1998.
- BUZZONI, Marco. Esternalismo/Internalismo. In: *Enciclopedia Filosofica Bompiani*, Milano: RCS Libri, 2006. p. 3702.
- CARTER, J. Adam; KALLESTRUP, Jeseper; PALERMOS, S. Orestis; PRITCHARD, Duncan. Varieties of externalism. *Philosophical Issues*, v. 24, pp. 63–109, 2004.
- CHALMERS, David J. Foreword,. In: CLARK, Andy. *Supersizing the mind: embodiment, action, and cognitive extension*, Oxford: Oxford University Press, 2011. pp. ix-xvi.
- CLARK, Andy. *Being there. Putting brain, body, and world together again*. The MIT Press, 1997.
- _____. Reason, robots and the extended mind. *Mind & Language*, v. 16, n. 2, pp. 121-145, 2001.
- _____. Memento's revenge: the extended mind extended. In: MENARY, Richard (ed.). *The extended mind*. Cambridge: MIT Press, 2010a. pp. 43-66.
- _____. Coupling, constitution and the cognitive kind. In: MENARY, Richard (ed.). *The extended mind*. Cambridge: MIT Press, 2010b. pp. 81-99.
- _____. *Supersizing the mind: embodiment, action, and cognitive extension*, Oxford: Oxford University Press, 2011.
- _____; CHALMERS, David. J. The extended mind. *Analysis*, v. 58, n. 1, pp. 7-19, 1998. Reprinted in MENARY, Richard (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 27-42.
- DENNETT, Daniel C. *The intentional stance*. Cambridge, MA: MIT Press, 1987.
- _____. The myth of original intentionality. In: MOHYELDIN SAID, K. A.; NEWTON-SMITH, W. H.; VIALE, R.; WILKES, K. V. (Eds.). *Modelling the mind*. Oxford, England: Oxford University Press, 1990. pp. 43-62.
- _____. *Consciousness explained*. Boston. New York, London: Little, Brown and Co., Back Bay Books, 1991.
- DONALD, Merlin. *Origins of the modern mind. Three stages in the evolution of culture and cognition*. Cambridge University Press, 1991.
- GOLDIN-MEADOW, Susan. *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press, 2003.
- GRUSH, Rick. In defense of some cartesian assumptions concerning the brain and its operations. *Biology and Philosophy*, v. 18, n. 1, pp. 53-92, 2003.

- HEERSMINK, Richard. Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences*, v. 14, n. 3, pp. 577-598, 2015.
- HEIDEGGER, Martin. Being and time. Translated by Joan Stambaugh. SUNY Press, 1996.
- HURLEY, Susan. Varieties of externalism. In: MENARY, Richard (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 101-153.
- KENNY, Anthony. *The metaphysics of mind*. New York/Oxford: Oxford University Press, 1992.
- KIVERSTEIN, Julian; FARINA, Mirko. Embraining culture: leaky minds and spongy brains. *Teorema: Revista Internacional De Filosofía*, v. 30, n. 2, pp. 35-53, 2011.
- LASHLEY, Karl S. In search of the engram. In: F. BEACH; D. O. HEBB; C. MORGAN; H. NISSEN (eds.). *The neuropsychology of Lashley*. New York: McGraw-Hill, 1950.
- MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010a.
- _____. The holy grail of cognitivism: a response to Adams and Aizawa. *Phenomenology and Cognitive Sciences*, v. 9, n. 4, pp. 605-618, 2010b.
- MERLEAU-PONTY, Maurice. *Phenomenology of perception*. Oxford: Routledge, 2012.
- PALERMO, S. Orestis. Loops, constitution, and cognitive extension. *Cognitive Systems Research*, v. 27, pp. 25-41, 2014.
- PRESTON, John. The extended mind, the concept of belief, and epistemic credit. In: MENARY, Richard (ed.). *The Extended mind*. Cambridge: MIT Press, 2010. pp. 359-362.
- PUTNAM, Hilary. The meaning of meaning. *Minnesota Studies in the Philosophy of Science*, v. 7, pp. 131-193, 1976.
- QUINE, Willard van Orman. Natural kinds. In: *Ontological Reality & Other Essays*. New York: Columbia University Press, 1969.
- RORTY, Richard. *Philosophy and the mirror of nature*. Oxford: Basil Blackwell Ltd, 1980.
- ROSS, Don; LADYMAN, James. The alleged coupling-constitution fallacy and the mature science. In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 155-166.
- ROWLANDS, Mark. Consciousness, broadly construed. In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 271-294.
- RUPERT, Robert. D. Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, v. 101, n. 8, pp. 389-428, 2004.
- _____. Representation in extended cognitive systems: does the scaffolding of language extend the mind? In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 326-328.

SCHMITT, Richard. Heidegger's analysis of 'tool'. *The Monist*, v. 49, n. 1, pp. 70-86, 1965.

SIMONS, Daniel J.; RENSINK, Ronald A. Change blindness: past, present and future. *Trends in Cognitive Sciences*, v. 1, n. 7, pp. 261-267, 2005.

STERELNY, Kim. Externalism, epistemic artifacts and the extended mind. In: SCHANTZ, Richard (ed.). *The externalist challenge. New studies on Cognition and Intentionality*. Berlin & New York: de Gruyter, 2004. pp. 239-254.

_____. Minds: extended or scaffolded? *Phenomenology and Cognitive Sciences*, v. 9, pp. 465-482, 2010.

SUTTON, John. Exograms and interdisciplinarity: history, the extended mind, and the civilizing. In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 189-225.

WHEELER, Michael. In defense of extended functionalism. In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 245-270.

WILSON, Robert, A. Meaning making and the mind of the externalist. In: MENARY, Richard A. (ed.). *The extended mind*. Cambridge: MIT Press, 2010. pp. 167-188.

Recebido em: 03-03-2019

Aceito para publicação em: 25-06-19