

The Self Between Vehicle Externalism and the Myth of the Cartesian Theatre¹

Giuseppe Vicari

RESUMEN

Este artículo discute el dilema propuesto por Andy Clark, según el cual o bien aceptamos los argumentos de la paridad para concluir que el yo puede estar constituido parcialmente por estructuras del entorno, o bien nos comprometemos implícitamente con el mito del teatro cartesiano. En primer lugar, objeto que los argumentos de la paridad de Clark asimilan las relaciones de nivel personal entre organismo y estructuras del entorno, mediadas por la conciencia y la intencionalidad, a las relaciones causales auto-organizativas que median en la incorporación y el mantenimiento de recursos internos. A continuación, critico la afirmación de Clark de que el internismo se compromete implícitamente con el mito del teatro cartesiano, pues es así sólo en la medida en que el internismo comete el error categorial de adscribir propiedades de los contenidos personales a los vehículos subpersonales. Hay una tercera vía entre el externismo de los vehículos y el mito del teatro cartesiano, según la cual nos vemos a nosotros mismos como organismos configurados de manera variable por nuestra situación en el entorno así como por nuestro cuerpo biológico, sin por ello recolocar al sujeto mismo en las estructuras del entorno que le configuran.

PALABRAS CLAVE: *conciencia, intencionalidad, auto-organización, mente extendida, argumentos de la paridad, trasfondo.*

ABSTRACT

This paper targets Andy Clark's dilemmatic argumentative structure, according to which either we accept the parity arguments for the conclusion that the self can be in part constituted by environmental structures, or we are implicitly committed to the Myth of the Cartesian Theatre. First, I criticize Clark's parity arguments because they assimilate the personal level relationships between organism and environmental structures, mediated by consciousness and intentionality, to the self-organizing causal relationships mediating the recruitment and maintenance of internal resources. Then, I criticize Clark's claim that internalism is implicitly committed to the Myth of the Cartesian Theatre because this argument works only insofar the internalist makes the category mistake of ascribing properties of personal level mental contents to subpersonal vehicles. A third view between vehicle externalism and the Myth of the Cartesian Theatre is available: seeing ourselves as organisms variously shaped by our

ecological embedment and biological embodiment without relocating the subject itself in the environmental structures shaping it.

KEYWORDS: *Consciousness, Intentionality, Self-Organization, Extended Mind, Parity Arguments, Background.*

INTRODUCTION

Non-standard realizations of psychological properties are quite common in functionalistic literature. However, when compared, for example, with Daniel Dennett's science fiction story (1978), in which Dennett's body and point of view are located in Tulsa while his brain is kept alive in Houston, the theory of the extended mind (EM) has at least two new features:

1. Extended, biotechnological agents exist right here and now. To realize an environmentally distributed agent is sufficient to have low-tech but transparent, well-integrated equipment such as pen and paper or a wristwatch.
2. While Dennett's fiction apparently does not deny that the brain is the ultimate (though remote) controller of behavior, and is, therefore, compatible with internalism [Bartlett (2008)], the main idea illustrated by EM is that environmental and internal structures jointly govern behavior. In this sense, Clark says, we are not mere "co-oons" of the real, organic self. Rather, "We are 'soft-selves', continuously open to change and driven to leak through the confines of skin and skull, annexing more and more nonbiological elements as aspects of the machinery of mind itself" [(2007), p. 112].

To argue for these points, Clark uses two strategies. On one side, he argues that the internal structure of the organism is self-organized and loosely coupled like the structure of an extended coalition. So, by parity of reasoning, there is no reason to identify the organism as the only individual bearer of cognition and agency. On the other side, Clark argues that internalism is necessarily committed to the Myth of the Cartesian Theatre, because it implies the identification of self and consciousness.

I will argue that both arguments fail. The first argument fails because it assimilates personal level, organism-environment relationships mediated by consciousness and intentionality to the internal causal dynamics, thereby losing sight of the properly active role played by the organism in extended coalitions.

The second argument fails because internalism involves the Myth of the Cartesian Theatre only if we accept the category mistake of attributing properties of contents to subpersonal vehicles.

I. OTTO'S THOUGHT EXPERIMENT

Since Clark and David J. Chalmers' 1998 seminal paper "The Extended Mind," vehicle externalism about mind and cognition has been seen as implying vehicle externalism about the cognizing and acting self.² In that paper, the main argument for vehicle externalism is a now-famous thought experiment. Otto suffers from Alzheimer's disease and uses a notebook instead of his compromised biomemory. Like his healthy friend Inga, Otto learns that there is an exhibition at MOMA and desires to go there: Otto looks at the notebook, retrieves the information that MOMA is on 53rd Street and goes there, while Inga desires to go, remembers that MOMA is on 53rd Street and goes there [pp. 12-3].

Now, according to EM, Otto and Inga's cases parallel: Otto has a dispositional belief about the address even before looking at the notebook, and Inga has the same belief even before retrieving it from biomemory. The cases are behaviorally distinguishable but they exhibit sufficient functional similarity, namely they have "*enough of the same central features and dynamics of a normal agent having ... the dispositional belief*" [Clark (2005), p. 7].

Clark and Chalmers analyze various commonsense requirements (portability, easy access, reliable access) that vehicles must satisfy to count as proper parts of the physical realization of a coarse-grained functional role. They define a set of criteria aimed, among other things, at avoiding unacceptable belief-attributions³:

1. The external support is a constant in the organism's life.
2. The information stored in it is directly available without difficulties.
3. The information is automatically endorsed to carry out the process.
4. The information has been consciously endorsed in the past [Clark and Chalmers (1998), p. 17].

Now, the point of the thought experiment is that the two cases are functionally and explanatorily on a par: so, since what counts for an information to be part of a cognitive or mental process is its causal role, if external sources contain information poised to play that role they count as vehicles of mental contents. "Otto *himself*", then, would be better seen "as an extended system, a coupling of biological organism and external resources" [ibid., p. 18].

II. THE EXTENDED SELF

But why should we conceive of Otto-with-the-notebook as an individual self rather than saying that Otto, the cognitive and agentive self, is using a notebook to achieve his goals? Is mere coupling sufficient to extend the self? As Terry Dartnall would say, if I dig a hole with a spade it's me, and not me-and-the-spade, who gets the prize for the best hole in the garden even though I could not dig the hole without the spade [(2004), p. 145].

The argument running behind Dartnall's intuition is that EM commits a causal-constitutive fallacy [Adams and Aizawa (2001); (2009)]: the fact that a system *X* is causally dependent on an external source or object *Y* to perform its operations does not imply that *Y* is a constitutive part of *X*. Dartnall's example is meant to illustrate, then, that blurring the boundaries between the causal and the constitutive implies, among other things, losing sight of the subject as it is traditionally conceived, that is as a locus of experience, rationality, and agency, capable of robust performance across different environments and tasks [Rupert (2004)].

In fact we could notice, with Rob Rupert, that the robust performance of cognitive systems are explained by the fact that internal processes constitute the set of cognitive, agentive, and experiential capacities that we call "the self." This does not imply that we never use tools to achieve cognitive and agentive goals, or that we never structure the environment so as to transform problem-solving via epistemic actions.⁴ Indeed, *we* use tools and *we* perform actions, both pragmatic and epistemic.⁵ On the contrary, the claim that psychological properties supervene on contingent, temporary, and shifting biotechnological coalitions whose components "play an active causal role, and ... jointly govern behaviour" [Clark and Chalmers (1998), p. 8] would obscure the properly active role played by the agent in organizing and managing the complex web of environmental resources used to achieve our goals.

Clark's answer to this objection is that the "agent-as-the-tool-user" picture obscures the fact that the subject as traditionally conceived is itself nothing but a set of contingently coupled resources, self-organizing their contribution to cognition and action without the intervention of a "central controller" micromanaging the fine-grained details of action and cognition:

There is *no self* if by self we mean some central cognitive essence that makes me who and what I am. In its place there is just the "soft-self": a rough-and-tumble control-sharing coalition of processes – some neural, some bodily, some technological – and an ongoing drive to tell a story ... in which 'I' am the central player. [(2007), p. 114]

The massive modularity of the neural architecture of the visual cortex is especially interesting [Milner and Goodale (2007); cf. Clark (2003), pp. 99ff.;

Clark (2007), pp. 109-10]: conscious perception, in fact, does not micromanage visuomotor behavior. Rather, the kind of control involved in visuomotor behavior is similar to the functioning of telerobotic systems, where an “intelligent operator” sets the high-level goals of the entire system (“grasp *X*”), while a robot-slave performs and controls the fine-grained details of movements.

In the same way, the organization of visuomotor behavior emerges from two separate pathways in the brain: the “ventral stream” with its reciprocal connections with the “cognitive areas,” and the “dorsal stream,” connected with the motor areas.

The fine-grained details of movements (like grip aperture) are under control of the ventral stream, which operates outside the scope of visual consciousness: in fact, visuomotor behavior is almost intact in patients suffering from visual form agnosia caused by impairment of the ventral streams [Goodale et al. (1991)]. Also, conscious perception, unlike visuomotor behavior, is deceived by Titchener’s illusion: two identical circles are seen as different in size due to Gestalt effects, so that we would expect different grip apertures when subjects grasp the two circles. However when subjects are asked to grasp the circles, their grip aperture is identical in both cases [Aglioti et al. (1995)].

According to this model, the conscious self is not, then, the author of the action—at best, it is its “business manager” [Clark (2007), p. 110]: it sets high-level goals but it does not control the action because the performance of the actual movements exploits the motor competences of other subsystems. The conscious self involved in visuomotor behavior is an “open-ended ecological controller” [ibid., p. 103], a mechanism adding some “crucial nudges” (conceptual analysis, deliberative reasoning and thought) to a mix of different resources organized by loose causal relationships, exploiting the passive and self-organizing dynamics exemplified by Rodney Brooks’ “subsumption architectures” [Brooks 1991].

The case of “thoughtful gestures” provides another example of self-organizing, distributed cognition without central control. Susan Goldwin-Meadow [(1999); Goldin-Meadow and Wagner (2005)], Shaun Gallagher [(2005), chap. 5], and others argue that some cases of gesturing are not simply expressions of ready-made thoughts in the head. These gestures do not communicate anything to other people. In fact, blind people gesture when talking to other blind people. Moreover, though they are not fully automatic, these gestures are largely out of conscious control. Now, following Goldin-Meadow, Clark argues that gestures are a part of “an integrated ‘language-thought-hand system’” [(2008), pp. 123ff.]. For example, gesture-speech mismatches in the explanation of a (wrong) solution to a problem “tell” the system that different ideas, realized in different representational formats, are competing. Mismatches tell that the system is at a transition point – and in fact, mismatches are faster in learning the right solution, but they are more

likely to lose the ability to apply it without further training [Goldin-Meadow and Wagner (2005), p. 236].

In these cases, distributed and loosely coupled subsystems, self-organized in feedback loops, all play an active role in problem-solving activity: they compete by representing different ideas, and there is no “central controller” that adjudicates the dispute. Starting from here, it takes just a “slippery slope” move to substitute or complement gestures with pen and paper. Using marks on paper in the process of thought-making can play a cognitive role exactly for the same reasons that gestures do: written language presents ideas and ongoing reasoning encoded in a different representational format that competes with other ongoing thought processes. Moreover, the externalization of the process of thought-making, the “freezing” of thoughts on paper, lightens the computational load in the head and helps in revising the steps of reasoning through interlocking perceptuomotor cycles [Clark (2008), p. 126; cf. Clark (2003), pp. 69-75].

Visuomotor behavior and the cognitive role of gestures illustrate, then, that agency and cognition, as performed by the biological organism, are the result of a loosely coupled coalition of mixed resources with no central controller. Already at the internal level, cognition and action involve the contingent, soft assembly of mixed resources with no “intelligent planning,” conscious choice, or monitoring: in cases like gestures, there is no conscious choice and control, but just “complex, subpersonally integrated routines that are selected for their peculiarly cognitive virtues.” In such cases, where conscious choice is missing, “the choice consists only in the emergence of an effective distributed problem-solving whole” [Clark (2008), p. 137]. The same holds true at the macrolevel: when a cognitive problem is solved by a complex system made out of an organism reliably coupled with computer, pen, paper, and the like, and all the elements play an active causal role in the problem solving activity, then the distribution of cognitive load demands the spread of cognitive credit [Clark and Chalmers (1998), p. 8].

The illusion of a “central” self micromanaging the actual working of the self-organizing coalition would be produced under the concomitant pressure of two different forces [Clark (2007), pp. 114-5]: on one side, society needs to reject certain patterns of behavior, which requires the identification of an author of the patterns in question even though, from a scientific point of view, the patterns themselves are just the result of the cooperative effort of distributed mechanisms.

On the other side, as Wegner (2002) has argued, “the illusion of conscious will” (that is, the illusion of being an agent whose will causes actions) has a functional role to play: it informs the organism with an “emotion of authorship” that allows self-ascription of some of the events happening in the world (the actions of the organism). Also, using Dennett’s concept of a “narrative self,” we can say that once we discover the simplifying power of a co-

herent narrative centered on “the self,” we tend to protect the narrative itself by identifying a persisting protagonist of it: the conscious self becomes the (putative) author rather than a product of the narrative [Dennett (1991), pp. 415-8].

The boundaries of the self, after all, are negotiable even from the first-person point of view. In the famous “rubber hand” experiment, a subject sees a dummy hand while one of his real hands is hidden. Both the unseen real hand and the seen dummy hand begin to receive tactile stimulation and, after a while, the subject feels tactile stimulation in the dummy hand [Botvinick and Cohen (1998); cf. Ramachandran and Blakeslee (1998)].

In extreme cases, using the same principles of the dummy hand experiment, the sense of self can extend entirely outside the body. In a set of experiments produced in virtual reality scenarios, people have genuine “out-of-body experiences”: they experience a virtual projection of their own body in front of them as their real body, and they experience their own point of view as located outside it [Ehrsson (2007); Lenggenhager et al. (2007); Metzinger (2009), chap. 3].

These experiments would show, then, that the body image is “a mental construct, open to renewal and reconfiguration” [Clark (2003), p. 61]. Likewise, according to Clark, our very notions of self, person, and the relative attribution of agency and responsibility are a “*forensic matter*,” that is, “a matter of legal and moral convenience more than metaphysics, and a convenience, moreover, that must become increasingly inconvenient as science and technology progress” [Clark (2005, p. 10)].⁶

III. THE PRIORITY OF THE CONSCIOUS SELF IN EXTENDED COALITIONS

Though I agree with various aspects of Clark’s analysis – the emphasis on the desubstantialization of the self, the key role of perceptuomotor loops and in general of biological embodiment, and environmental embedding for the structure of the self – I do not think that parity arguments are sufficient to establish the radical conclusion that Otto-and-the-notebook constitute a single, extended self.⁷

The critical point in Clark’s analysis is the slippery-slope move from examples of biological self-organization to the claim that the same mechanism is working when the organic coalition “incorporates” aspects of the local environment through epistemic actions:

The relations between our conscious sense of self ... and the many nonconscious neural goings-on that structure and inform this cognitive profile are, it seems to me, pretty much on a par with the relations between our conscious minds and various kinds of transparent, reliable, robust, and readily accessed nonbiological resources. When those resources are of a recognizable knowl-

edge- and information-based kind, the upshot is an extended cognitive system: a biotechnologically hybrid mind, a biotechnologically hybrid self" [(2007), p. 108].

The recruitment and maintenance of a distributed coalition are underlain by the same mechanisms in the internal and external cases. There is no conscious recruitment and control of gestures. Rather, there is just the self-organizing emergence of a distributed coalition "similar" to cases "when we are busy thinking and writing at the same time" [Clark (2008), p. 126]. These are cases that cannot be captured by the model of a "skull-bound intelligent agent" consciously offloading part of the cognitive work. Otto himself, according to Clark and Chalmers' description, does not consciously offload information in the notebook; rather, his use of the notebook is "so well-practiced as to become automatic and unreflective" [ibid., p. 245].

This description, I argue, leaves out the specific roles played by consciousness and intentionality in mediating the interactions between the organism as a whole and the environment.

Consider the criterion of conscious past endorsement. According to this criterion, information stored in an external device "has been consciously endorsed at some point in the past and indeed is there as a consequence of this endorsement" [Clark and Chalmers (1998), p. 17].

Conscious past-endorsement is the most controversial among the criteria of reliable coupling because, as Clark and Chalmers note [ibid.], it seems plausible to say that information unconsciously endorsed can drive behavior – think about hypnosis and subliminal perception. However, as Rupert noted, the criterion is necessary for EM because its elimination would produce highly implausible, and perhaps even inconsistent, attributions of beliefs.

Giving up conscious past endorsement would imply, for example, that a heavily internet-dependent subject believes all the information made available by Google just because the other conditions for reliable coupling are satisfied. But if conscious past endorsement holds, then external structures can play a cognitive role only insofar as the conscious self mediates the relationships between organism and environment, and the tool-use interpretation of Otto's thought experiment would look more plausible to capture this fact [Rupert (2004), pp. 401-5].

But perhaps this is not the right way of characterizing the consequences of conscious past endorsement: Clark could reply that the conscious self is playing the role of an ecological controller, which first endorses information and then sets high level goals and plans for the operation of separated subsystems without micromanaging their work.

I think, however, that Rupert's point marks a difference between internal dynamics as described by Clark, where temporary coalitions emerge from self-organization, and extended coalitions, where external items are recruited

and maintained by the organism. Given conscious past endorsement, then, the organism seems to be the only properly active part of the temporary, biotechnological, extended coalition of resources. Rupert, however, thinks that this point can just soften the externalists' rhetoric given their deflationary view of the role and nature of consciousness. On the contrary, I think that the objection from the central role of the conscious organism can be generalized to other central features of cognition and agency, such as the holism of intentionality and the attribution of authorship of actions.

Daniel Weiskopf has challenged EM on the basis that reliable coupling is not sufficient to account for holism (2008). Beliefs and intentional states in general are partly constituted by the network of other states in which they are placed, where the network is regulated by normative criteria of rational assessment.⁸

It is part of the functional role of a belief that its content is subject to informational integration and updating: so Inga's belief that MOMA is on 53rd Street will be updated on learning that the museum has been torn down, and other relevant beliefs will change accordingly: if Inga has the belief that the museum cafe makes a good latte she will, for example, modify that belief to the past tense.

Does the notebook satisfy this criterion? Weiskopf's thesis is that since "most alleged cases of externally located mental states do not share this feature [then], by the functionalist principle, they cannot be beliefs" [(2008), p. 268]. Since the notebook, unlike Inga's biomemory, will not automatically erase or modify relevant beliefs in light of new and relevant information, we would be forced to ascribe to the extended system a degree of irrationality that we would hardly attribute to Inga. In addition, if a system stores patently contradictory information, then there will be no satisfactory mentalistic explanation of his behavior and EM's explanatory power will be seriously threatened.

Of course, Otto could check every page of the notebook when he stores or retrieves information. But, beyond Weiskopf's argument, when we accept this point, we get an interesting result: not only consciousness is required for past-endorsement, it is also required to maintain the holistic unity of the extended mind. Indeed, it is difficult to see how even a simple notebook could achieve a transparency comparable to internal biomemory: if Otto must consciously check every page looking for inconsistencies, how could he use the notebook automatically and unreflectively?

So, while the standard internal case does not require conscious control, Otto's case requires it by default, so that the description of Otto as the only active agent in the putatively extended Otto-and-the-notebook coalition looks more plausible than Clark's description.

What should be said about more "active," semi-autonomous devices that can run some processes on their own, such as alarm clocks, computers and robots?

Suppose now, following Brie Gertler (2007), that Otto is reliably coupled to a high-tech notebook in which he stores information that can work as dispositional beliefs, desires, and intentions: he writes that he wants to cook a cake, that he needs the ingredients, that he needs to go to the store to buy them, that the store is located in such and such a place, and so on. By the criteria of reliable coupling, the notebook works as a proper part of Otto. Suppose further, that the notebook is a proper part of an electronic brain linked to a robotic body, and that this set of devices is allowed to plan and perform actions based on the information stored in the notebook. The biological Otto works like the “ventral stream” described by Goodale and Milner, or like the “intelligent operator” of a telerobotic system: Otto sets high-level goals while the fine-grained details of action would be “delegated” to the “dorsal,” notebook-robot subsystems.

Now while the organic Otto is sleeping the robotic subsystem carries out the action plan: it goes to the store, buys the ingredients, bakes the cake. Can we describe this sequence of events as Otto’s actions? If Otto himself extends beyond the biological boundaries just because internal and external dynamics mirror each other, and if by “action” we mean movements caused in the appropriate way by mental states, then the answer is “yes.” Not only Otto is responsible for the action:⁹ according to EM, he is the actual author of it even when he sleeps, just because Otto and the external resources satisfy the reliability criteria and Otto sets the goals of the system.

Perhaps Clark could reject this implausible consequence because, since Otto is sleeping, there is no conscious control.¹⁰ But even in this case, then online, direct, conscious control seems necessary to describe the extended coalition and the putative extended self as “acting.” So the organic Otto seems more accurately described as the tool user and “recruiter” of the extended coalition rather than as a mere subsystem of an extended biotechnological bearer of psychological properties.

IV CLARK’S SCIENTIFIC EVIDENCE FOR EM

If correct, the arguments described so far show that despite Clark’s emphasis on the parity between internal and external self-organizing processes, only the agent as traditionally conceived exhibits the kind of causally active contributions necessary to give rise to an extended problem-solving system. This is relatively uncontroversial: Clark himself admits that we cannot build a subject out of Otto-style notebooks [(2005), p. 6]. That’s why, as he says, cognition (but the same conclusion goes for the self) is organism-centered, but not organism-bound: it is the brain that recruits external structures, off-loads cognitive work and, thereby, transforms cognitive tasks [Clark (2008), pp. 122-3]. However, if we accept the claim that “what makes some informa-

tion count as a belief is the role it plays” [Clark and Chalmers (1998), p. 14], the idea of an organism-centered self looks like an ad-hoc move aimed at keeping track of the subject as traditionally conceived and of its active role.

Moreover, we must notice that Clark’s interpretation of the scientific evidence is controversial. First, in cases like the dummy hand and the out-of-body experiments, experience is not self-validating about its own subpersonal vehicles: that is, the fact that my experience includes the virtual body or the dummy hand as “mine” does not imply that the dummy hand or the virtual body are now external vehicles of experience. On the contrary, Clark himself claims that the experiments show that our body image depends on perceived correlations between bodily movements and sensory effects. In other words, what happens in these cases is that visual feedback “dominates” proprioceptive feedback: the brain makes proprioception coherent with vision in the same way that it makes hearing and vision coherent when we hear voices coming from the mouths of movie characters in the large screen of the movie theatre rather than from side-placed acoustic devices. The explanation of these phenomena is, then, internalistic: the dummy hand or the virtual body plays no properly active causal role in explaining self-experience.¹¹ In fact, even Clark writes that “Experience is, of course, *no more* than a clue. I do not mean, here or elsewhere, to advance any argument of the form ‘it seems to us as if we are/are not cognitively extended; therefore we are/are not extended’!” [(2008), p. 238n8].

Second, as Jerry Fodor (2009) says, there is no valid slippery-slope argument going from internal dynamics to vehicle externalism: even if we accept that gestures themselves rather than their neural bases play a cognitive role, more arguments are needed to show that coupling pen and paper to an organism during a problem-solving task relocates the external sources as constitutive parts of the agent. The possibility that an agent is simply using pen and paper still looks a more natural interpretation of the facts. The same goes for the comparison between advanced telerobotic systems and the functioning of the visual cortex. In general, one cannot validly infer the identity of concept *A* and concept *B* from the fact that the difference between application of concept *A* and application of concept *B* is a matter of degree.

V. TOOL-USE AND THE MYTH OF THE CARTESIAN THEATRE

If we can doubt the logical validity of parity arguments and EM’s interpretation of the empirical data is controversial, it is now clear why Clark claims that the main theoretical payoff of EM is the overcoming of the Myth of the Cartesian Theatre, which implicitly shapes the view that external items are tools used by the organism [Clark (2007), p. 109; (2008), pp. 136-9].

Clark's argument goes as follows: if we look at Otto's notebook as a tool used by Otto, then we will have to apply the same reasoning not only to Otto's biological arm, but also to Inga's biomemory. Following this pattern of reasoning, the biomemory would be the tool used by, for example, the frontal cortex and the dorsal stream could be a tool used by the ventral stream, which is, in turn, used by some other subsystem and so on, until the regress stops at some particular brain area – the Cartesian Theatre – where “I,” the conscious self, have my seat, issue commands, inspect all the information, and have the final say.

But one cannot reduce the self to the physical machinery of consciousness without losing sight of the psychological continuity and cohesion of the subject, as Libet's worries about free will have shown. Libet's studies of voluntary actions are taken to show that the brain prepares some movements before “I” have chosen to perform them: so it looks as if I am out of the loop simply because it is the unconscious brain, and not me, who has the “final say” [see Clark (2007), p. 113].

As Dennett (2003) rightly sees, the standard interpretation of Libet derives from a confusion between properties of contents at the personal level and properties of the explanatory vehicles at the subpersonal level: for example, from the fact that it seems to us as if we decide in a single instant what to do, it doesn't follow that the subpersonal mechanisms of the decision must also operate in a single instant. Rather, since the vehicles of decision making are spatially distributed in the brain, information processing relevant to the task will require time, despite the phenomenological impression of deciding “at time *t*.” If we give up the idea that mental processing requires a single vehicle and operates in a single instant, then the right conclusion is not that I am out of the loop, but rather that I am the loop.

In other words, the traditional view of the self as the organism-bound-tool-user is committed to the identification of self and mind with consciousness and, therefore, to a systematic misdescription of the relationships linking the conscious mind to the unconscious processes that shape and partly constitute our cognitive profiles. But, of course, commonsense would refuse to identify the self with any single brain area – likewise, it would refuse to eliminate the self on the grounds of the evidence of the existence of distributed coalitions with shared control of action and cognition.

Apparently, then, we are left with a dilemma: either we recognize unconscious processes and structures as partly constitutive of who and what we are, thereby opening the door to EM, or we shrink mind and self to the machinery of consciousness, thereby implicitly accepting the Myth of the Cartesian Theatre [Clark (2007), p. 108]. The arguments presented in the previous sections show that, if correct, the first horn of the dilemma describes a false implication: that is, even though we recognize (as we ought to do, for the reasons correctly stated by Clark and Dennett) unconscious neural processing

and internal self-organizing causal dynamics as partly constitutive of the self, it doesn't follow that these dynamics are on a par with relationships between organism and environment because relationships at the personal level are mediated by consciousness and intentionality.

Clark would, of course, reply that this characterization of the organism-environment relationship at the personal level "depicts ... outer resources as doing their work only by parading structure and information in front of some thoughtful internal overseer" [(2008), p. 137]. But this reply applies to EM too: conscious past-endorsement, the holistic structure of intentionality, and the notion of responsible agency in extended systems require conscious monitoring – that is, they require that the ongoing activity of the external structures be monitored by an "inner overseer," the conscious organism itself.

Does the traditional notion of the subject as captured by the agent-as-the-tool-user interpretation of Otto's case require something similar? Of course, Brie Gertler (2007) argues that the only way to block the parity argument and its implausible consequences is to endorse a "narrow view of mind" that shrinks the boundaries of the mind to the vehicles of consciousness. But this radical move (correctly criticized by Clark) is not the only theoretical alternative to EM.

The reason why internalism is not necessarily committed to the Myth of the Cartesian Theatre can be stated using the distinction between vehicles and contents. Recall that Clark's argument states that the agent-as-the-tool-user view seems "sensible and proper Until we turn the magnification on the biological brain itself" [Clark (2007), p. 111]. But why should we say that seeing Otto as a tool user at the personal level commits us to "turn the magnification" at the subpersonal level? It seems possible to say that there is real tool-use at the personal level while denying that there is any tool use at the subpersonal level. Indeed, given the distinction between contents and vehicles, seeing tool use at the subpersonal level looks like a category mistake: it makes no sense to say that the ventral stream uses the dorsal stream for the same reason that it makes no sense to say that the dorsal stream acts and the ventral stream perceives. It is the agent that sees and acts, and these capacities at the agent level are explained at the subpersonal level.

VI. CONCLUSION: SOME NOTES FOR A THIRD VIEW BETWEEN VEHICLE EXTERNALISM AND THE MYTH OF THE CARTESIAN THEATRE

Clark, then, has put forward a false dilemma: a third option is open to the internalist, that is, to use Clark's words, seeing the acting and cognizing self as the "cumulative effect of the coactive unfolding" of embodied and embedded conscious and unconscious processes, while refusing to blur the boundaries between self and world, user and tool. One of the main points of

interest of EM is that it tries to show that non-standard realizations of minds are not only logically possible, but already exist in the real world. Up to now, however, I do not see any convincing argument for the conclusion that the causal basis of selfhood extends beyond the boundaries of the organism. Clark has admirably conveyed the richness of the results of cognitive science and of its implications for our self-understanding as embodied creatures deeply embedded in the real world. Still, being contextualized creatures is not the same as being contingent coalitions distributed across brain, body, and world: one can fruitfully study the subtle and surprising ways in which body and world shape the mind without relocating the mind itself outside the embodied creatures that we are, and without assimilating personal, intentional relationships with subpersonal causal dynamics.

In John Searle's perspective, mental states are caused by and realized in the brain [cf. Searle (2004)]. This means that, in a sense, brain states are all we have for the purpose of representing the world [Searle (1983), p. 230]. Does this mean that we somehow are inside our heads? According to Searle, the answer is a definite "no": tools for representing the world and interacting with it work only insofar as they somehow "impact ... on our nervous systems." Nonetheless we do not, so to speak "crawl inside our own heads," rather, "we are *identical* with our own bodies, and the part of our bodies most important for our Intentional relations with the world is inside our skulls" [Searle (1991), p. 240].

Internalism by itself, then, does not imply that we are "somehow hidden within the machinery whose operations are most accessible to consciousness" [Clark (2007), p. 109]. Nor is internalism forced to claim that consciousness and intentionality describe the full range of the organism-environment relationships. Indeed, Searle and Clark could also agree on the fact that a bundle of "taken for granted skills, knowledge, and abilities ... structures and informs our sense of who we are, what we know, and what we can do" [Clark (2007), p. 106].

To use Francesca Di Lorenzo Ajello's words, the real innovative point of Searle's perspective can be found in its developing an account showing that intentionality has its roots "in a common and intersubjective Background of nonrepresentational practices and skills" [Di Lorenzo Ajello (2001), p. 72, and chap. 3 *passim*]. In this way Searle's internalistic account, "far from denying [...] intersubjectivity," as Putnam, Apel and Habermas claim, can systematically account for it in the general framework of a logical analysis of intentional states.

That is, although the brain is the causal basis of mental phenomena and selfhood; even a simple intention such as "to go to the refrigerator and get a bottle of beer to drink" requires the possession and exercise of biological and cultural skills such as "standing, walking, opening doors, pouring and drinking" [Searle (1983), p. 143] where there is no "sharp dividing line" between

skills regarding “how things are” and “how I do things” [ibid., p. 144]. The possession of these skills derives from the fact that “each one of us is a biological and social being in a world of other biological and social beings, surrounded by artifacts and natural objects,” so that “the Background is indeed derived from the entire congeries of relations which each biological-social being has to the world around itself.” [ibid., p. 154].

Nonetheless, Background skills, realized “in human brains and bodies” are not a set of things in the world: “all this embeddedness is only relevant to the production of the Background only because of the effects that it has on me, specifically the effects that it has on my mind-brain. The world is relevant to my Background only because of my interaction with the world” (ibid.).

The self, in this perspective, is not an internal substance: rather, it is the cognizing and acting organism in its ecological interactions with the physical and social environment. As our sense of self grows and develops over time, language complements the structures of prelinguistic intentionality and consciousness, providing the key move to create an institutional reality [Searle (2006)] and, with it, the characteristic dialectic between “universal constraints of language and deontic commitments” which constitutes the characteristic cognitive profile of human rationality [Di Lorenzo Ajello (2009)].

Embodiment and embedding, then, do not turn us into mere Humean, extended bundles: rather, at the right level of description – the personal level – we are biologically embodied and socially embedded, causally real presences in the real world [Vicari (2008a), chap. 5; cf. Vicari (2008b)].

Dipartimento FIERI-AGLAIA

Facoltà di Lettere e Filosofia

Università degli Studi di Palermo

Viale delle Scienze, Ed. 12, 90128 Palermo, Italy

E-mail: giuseppecicari@unipa.it

NOTES

¹ This paper presents some results of the ongoing research that I am pursuing at the University of Palermo (postdoc research grant MIUR 2009). As a member of the Workshop of Critical Social Philosophy at the University of Palermo, I have benefited from critical discussions of previous versions of this work with my supervisor, Prof. Francesca Di Lorenzo Ajello, and my colleagues Claudia Rosciglione and Giancarlo Zanet. Thanks also to Julian Kiverstein for our discussions of the extended mind during my time of study at the University of Edinburgh.

² A remarkable exception is Wilson [(2004), pp. 141-3].

³ Clark (1997) defines the mentioned criteria to answer the objection that, based on a purely causal-informational account of cognition, it would be easy to say that, if I

have an Encyclopedia in my room, then I can be ascribed all the information contained in it. Especially the last criterion (conscious past endorsement) seems necessary to avoid these unacceptable belief-attributions. However, it is not clear whether Clark and Chalmers' criteria can be independently justified or if they are just *ad hoc* procedures designed to rule out *by fiat* some potential counterexamples.

⁴ For the difference between pragmatic actions aimed at achieving a goal and episodic actions aimed at transforming the problem-solving see Kirsh and Maglio (1994).

⁵ As Rowlands (2009) notes, many processes are designed to work only in causal coupling with environmental features, but this is not enough to include external components as proper parts of the process. My ability to play the piano cannot be exercised if a piano is not available and it is, in this sense, causally dependent on the piano, but the ability itself is not extended to the piano. In the same way, even though the existence of a property requires a relational explanation – or, epistemically, even though the individuation of a property is relational –, this tells us nothing about the location of a property.

⁶ Another case of technological extension of the self is given by so-called brain-machine interfaces: mechanical arms or computer monitors linked to, for example, the damaged brain of patients suffering from locked-in syndrome, who are fully conscious but incapable of movement and communication. These machines collect electrical signals from the brain and translate them into actions caused by those signals. Taking for granted the usefulness of these devices, it remains unclear what meaning should we give to these technologies: some writers claim that they restore the patients' lost cognitive capacities and therefore constitute a part of the self [Fenton and Alpert (2008)], while others [e.g., Walter (2009)] see them as expressive means of the patients' intact abilities.

⁷ I neither agree, among other things, with Clark's description of conscious life as a sort of Humean unstructured bundle of ideas [cf. Clark (2007), p. 107], nor do I agree with the supposed implication from the desubstantialization of the self to the kind of instrumentalism that we can find in both Clark and Dennett [Clark (2007), p. 114]. Indeed, I do not see how a forensic view of the conscious self could cohere with the (though limited) causal role that Clark recognizes to the ecological controller. See Searle (1992) for a better analysis of the structure of consciousness, and Vicari [(2008a); (2008b)] for the analysis of some implications of Searle's theory of the unity of consciousness and of the self-referentiality of perception and action for a realist and antisubstantialist view of the self placed within the framework of Edelman's and Damasio's theories of consciousness.

⁸ This point is independent of the acceptance of both functionalism and externalism [cf. Searle (1983), pp. 19-20; Dennett (1987), pp. 13-36].

⁹ In the same way that, say, Hitler is responsible for the invasion of Russia even though he was never physically there.

¹⁰ This reply is suggested by Clark's discussion [(2003), pp. 100ff.] of the failure of telerobotics to convey a sense of "real telepresence," when compared with "teleoperators" – systems in which the intelligent operator controls the details of movements of distant robotic systems. This failure is mainly due, according to Clark, to a lacking "sense of potential intervention" [ibid., pp. 107-8] and sense of the unfolding of the details of action. There are various technologically possible ways to provide these features in extended coalitions, but again this is irrelevant to my point,

which concerns the properly active role played by the conscious self even with respect to semi-autonomous external structures, where mere ecological control does not seem sufficient to account for responsible agency.

¹¹ Thanks to Julian Kiverstein for critical discussions on this point.

REFERENCES

- ADAMS, F. and AIZAWA, K. (2001), 'The Bounds of Cognition,' *Philosophical Psychology*, vol. 14, pp. 43-6.
- ADAMS, F. and AIZAWA, K. (2009), 'Why the Mind Is Still in the Head,' in Robbins P. and Ayede, M. (eds.), *The Cambridge Handbook of Situated Cognition*, Cambridge, Cambridge University Press, pp. 78-95.
- AGLIOTI, S., DESOUZA, J. F. X., and GOODALE, M. A. (1995), 'Size-Contrast Illusions Deceive the Eye But Not the Hand,' *Current Biology*, vol. 5, pp. 679-85.
- BARTLETT, G. (2008), 'Whither Internalism? How Internalists Should Respond to the Extended Mind Hypothesis,' *Metaphilosophy*, vol. 39, pp. 163-84.
- BOTVINICK, M. and COHEN, J. (1998), 'Rubber Hands 'feel' touch that eyes see,' *Nature*, vol. 391, p. 756.
- BROOKS, R. (1991), 'Intelligence without Representation,' *Artificial Intelligence*, vol. 47, pp. 139-60.
- CLARK, A. (2003), *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*, Oxford, Oxford University Press.
- (2005), 'Intrinsic Content, Active Memory and the Extended Mind,' *Analysis*, vol. 65, pp. 1-11.
- (2007), 'Soft Selves and Ecological Control,' in Spurrett D., Ross D., Kincaid H. and Stephens L. (eds.), *Distributed Cognition and the Will: Individual Volition in Social Context*, Cambridge (Mass.), MIT Press, pp. 101-22.
- (2008), *Supersizing the Mind. Embodiment, Action, and Cognitive Extension*, Oxford, Oxford University Press.
- CLARK, A. and CHALMERS, D. J. (1998), 'The Extended Mind,' *Analysis*, vol. 58, pp. 10-23.
- DARTNALL, T. (2004), 'We Have Always Been...Cyborgs,' *Metascience*, vol. 13, pp. 139-48.
- DENNETT, D.C. (1978), 'Where Am I?,' in Dennett, D.C., *Brainstorms*, Cambridge (Mass.), MIT Press, pp. 310-23.
- (1987), *The Intentional Stance*, Cambridge (Mass.), MIT Press.
- (1991), *Consciousness Explained*, Boston, Little Brown & Company.
- (2003), *Freedom Evolves*, London, Penguin Books.
- DI LORENZO AJELLO, F. (2001), *Mente, azione e linguaggio nel pensiero di John R. Searle*, Milano, Franco Angeli.
- (2009), 'Vincoli universali del linguaggio e impegni deontici nella costruzione della realtà sociale: da Hegel a Searle,' in Ruggi L., Testa, I. (eds.), *Lo spazio sociale della ragione*, Milano, Mimesis, pp. 423-46.
- EHRSSON, H. H. (2007), 'The Experimental Induction of Out-of-Body Experiences,' *Science*, vol. 317, p. 1048.
- FENTON, A., and ALPERT, S. (2008), 'Extending Our View on Using BCIs for Locked-in Syndrome,' *Neuroethics*, vol. 1, pp. 119-32.
- FODOR, J. (2009), 'Where Is My Mind?', *London Review of Books*, vol. 31, pp. 13-5.

- GALLAGHER, S. (2005), *How the Body Shapes the Mind*, Oxford, Oxford University Press.
- GERTLER, B. (2007), 'Overextending the Mind?', in Gertler, B. and Shapiro, L. (eds.), *Arguing about the Mind*, London, Routledge, pp. 192-206.
- GOLDIN-MEADOW, S. (1999), 'The Role of Gesture in Communication and Thinking,' *Trends in Cognitive Science*, vol. 9, pp. 419-29.
- GOLDWIN-MEADOW, S. and WAGNER, S. M. (2005), 'How Our Hands Help Us Learn,' *Trends in Cognitive Science*, vol. 9, pp. 234-41.
- GOODALE, M. A., MILNER, A. D., JACKSON, L. S., and CAREY, D. P. (1991), 'A Neurological Dissociation between Perceiving Objects and Grasping Them,' *Nature*, vol. 349, pp. 154-56.
- KIRSCH, D. and MAGLIO, P. (1994), 'On Distinguishing Epistemic From Pragmatic Action,' *Cognitive Science*, vol. 18, pp. 513-49.
- LENGGENHAGER, B. et al. (2007), 'Video ergo sum: Manipulating Bodily Self-Consciousness,' *Science*, vol. 317, pp. 1096-99.
- METZINGER, T. (2009), *The Ego-Tunnel. The Science of the Mind and the Myth of the Self*, New York, Basic Books.
- MILNER, A. D., and GOODALE, M. A. (2007), *The Visual Brain in Action*, 2nd ed., Oxford, Oxford University Press.
- RAMACHANDRAN, V., and BLAKESLEE, S. (1998), *Phantoms in the Brain*, New York, William Morrow.
- ROWLANDS, M. (2009), 'Enactivism and the Extended Mind,' *Topoi*, vol. 28, pp. 53-62.
- RUPERT, R. (2004), 'Challenges to the Hypothesis of Extended Cognition,' *Journal of Philosophy*, vol. 101, pp. 389-428.
- SEARLE, J. R. (1983), *Intentionality: An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press.
- (1991), 'Response to McDowell,' in Lepore, E. and Van Gulick, R. (eds.), *John Searle and His Critics*, Oxford, Blackwell, pp. 237-41.
- (1992), *The Rediscovery of the Mind*, Cambridge (Mass), MIT Press.
- (2004), *Mind: A Brief Introduction*, New York, Oxford University Press.
- (2006), 'What is Language: Some Preliminary Remarks', in Abel, G. (ed.), *Kreativität. XX Deutscher Kongress für Philosophie*, Hamburg, Felix Meiner Verlag, pp. 1223-48.
- VICARI, G. (2008a), *Beyond Conceptual Dualism. Ontology of Consciousness, Mental Causation, and Holism in John R. Searle's Philosophy of Mind*, Amsterdam/New York, Rodopi.
- (2008b), 'Incorporazione non riduttiva e autoreferenzialità causale. Note per una concezione realista del sé nella filosofia della mente di John R. Searle e nella teoria neurobiologica della coscienza di A. Damasio,' in Di Lorenzo Ajello, F. (ed.), *Razionalità, verità e mente*, Milano, Mondadori, pp. 207-20.
- WALTER, S. (2009), 'Locked-in Syndrome, BCI, and a Confusion about Embodied, Embedded, Extended, and Enacted Cognition,' *Neuroethics*, vol. 3, pp. 61-72.
- WEGNER, D. (2002), *The Illusion of Conscious Will*, Cambridge (Mass.), MIT Press.
- WEISKOPF, D. (2008), 'Patrolling the Mind's Boundaries,' *Erkenntnis*, vol. 68, pp. 265-76.
- WILSON, R. A. (2004), *Boundaries of the Mind. The Individual in the Fragile Sciences*, Cambridge, Cambridge University Press.