DISSOLVING THE SELF: THE COGNITIVE TURN OF THE EXTENDED MIND THEORY

Léo Peruzzo Júnior[1]

Amanda Luiza Stroparo[2]

Abstract: In this article, we intend to show how the theory of the extended mind, particularly Andy Clark's arguments, can explain mental processes not as restrictive phenomena to the brain and endorse their connection to the body and the environment. Therefore, initially, we reconstruct the main materialist perspectives that have enclosed the self to the cranium; then, we indicate how the extended character of the mind escapes its natural limits and blends "shamelessly" into the world. We argue that external artifacts play an important role in guiding our actions, so that changes in the environment can cause changes in the behavior of the cognitive agent, what constitutes a constitutive dependence. In this way, the extended mind thesis challenges both traditional functionalism and externalism as it considers cognitive processes, on the one hand, to be relevant interactions of the individual with the environment and, on the other, intention-driven behaviors. Through the integration of biological bodies with artifacts or tools, we support a reading that dissolves the classic "explanatory gap" of cognitive sciences.

Keywords: Cognition. Self. Extended mind. Brainocentrism. Andy Clark.

INTRODUCTION

One of the most debated topics in cognitive science, philosophy of mind, neuroscience and psychology is, according to Shaun Gallagher (2013) and Richard Heersmink (2020), the "self" and its implications for certain

[1] Professor at the Philosophy Department, Pontifical Catholic University of Paraná and FAE Centro Universitário, Curitiba, PR – Brasil. ORCID: https://orcid.org/0000-0003-3084-5170. Email: leo.junior@pucpr.br.

[2] PhD Student in Philosophy at Pontifical Catholic University of Paraná, Curitiba, PR – Brasil. https://orcid.org/0000-0002-4774-3544. Email: als.stroparo@gmail.com.

http://dx.doi.org/10.1590/0101-3173.2023.v46n2.p193

perspectives of the mind. The question addressed by these sciences is to explain whether there is an agent behind human action and thought and, consequently, what would be the nature and limits of this agent. Thus, rather than demonstrating that there are phenomena or properties that we can refer to as "personal identity," we need to understand how a semantic dilemma could form something more substantial. Throughout this discussion, we first argue that the misunderstanding originating from most theories of mind is that they presuppose the existence of the concept and, then, analyze its constitution.

According to materialist theories of the mind, mental states are an attribute of matter or the physical world, be it any or all matter, or a special state that exists in the highly organized nervous system of higher animals and humankind. Thus, the claim, that mental events are identical to neural events, has become rather peculiar, especially if we assume that neurophysiological concepts are highly complicated processes of branching patterns of neuron discharges. Thus, which ontological status should we assign to mental events that have appeared in a world hitherto considered monistically physical? Moreover, if that is really the case, should we dissolve the concept of *self* and situate it as an extension in the world?

In the article "The Self and Its Brain", published in 1977, John Eccles and Karl Popper argued that the problem of the relationship between our bodies and our minds, especially the link between brain structures and mental dispositions, was an extremely difficult question. According to them, the self-conscious mind is an independent entity actively engaged in reading the multitude of active centers in the modules of the connecting areas of the dominant cerebral hemisphere. In other words, "[...] the self-conscious mind selects from these centers in accord with its attention and its interests and integrates its selection to give the unity of conscious experience from moment to moment. It also acts back on the neural centers." (POPPER; ECCLES, 1985, p. 355).

This argument, however, has yet another development: the self-conscious mind continues to play an interpretative role and superior control over neural events due to a bidirectional interaction at the interface between World 1 and World 2. What would this mean to cognitive science and the implications of a dualist-interactionist hypothesis? For the authors, the unity of conscious

experience does not come from a final synthesis in the neural machinery but from the integration of the self-conscious mind concerning what it reads from the immense diversity of neural activities between brain connections.

The primacy given to the self-conscious mind reveals that it would be responsible, on the one hand, for the unitary character of the experience and, on the other hand, for the relationship between the neural events in the connecting brain. Finally, the self-conscious mind would play an integrating role, giving a degree of correspondence between neural events and mental phenomena, but not of identity. The fact is that the unity of conscious experience is provided by the self-conscious mind, not through the neural machinery of the connecting areas of the cerebral hemisphere. A neurophysiological theory, therefore, should be able not only to explain the diversity of brain events, that allows for a global character unified conscious experience, but also how "[...] the self-conscious mind is developed in order to give this unity of the self in all of its conscious experiences and actions." (POPPER; ECCLES, 1985, p. 362). However, if there is a correlation between neural events and the unity of conscious experience, where, then, should we locate the *self*?

Patricia Churchland (2013), in *Touching a Nerve:* the self as brain, seeks to answer the above problem by assuming that neuroscience can reveal the physical mechanisms underlying psychological functions while acknowledging that the strategy can be extremely risky if one tries to explain the macrolevels (psychological properties) in terms of microlevels (properties of neural networks). In this sense, already in the first lines of her work, she states that "[...] my brain and I are inseparable. I am who I am because my brain is what it is. Even so, I often think about my brain in different terms from those ones I use when thinking about myself. I think about my brain as *that* and about myself as *me*." (CHURCHLAND, 2013, p.11). Therefore, to the extent that an empirical hypothesis is tested in a way that dissolves the higher-level capabilities of the human brain, we would also be doomed to reconstruct categories such as "memory," "attention," and "reasoning" and, consequently, rethink linguistic categories as self-sufficient components for understanding mental processes. More precisely, Churchland argues that explanatory power, coherence, and economy will favor the hypothesis that consciousness is simply a pattern of neuronal activity. Obviously, the

possibility of preventing the advance of the dualist argument from a neurophysiological perspective still seems to be far from the hypotheses that advocate non-reductionism for the concept of "self" and, consequently, for what we refer to as *conscious mental life*. Hence, Churchland writes (2013, p. 225):

> [...] the Neurobiology of Consciousness can be addressed through different strategies, each of which targets this question: What are the differences in the brain between conditions when we are conscious and those when we are not? Once those differences begin to be clarified, then the next question is this: What mechanisms support and regulate conscious states? With progress on those two questions, we may be able to address why hunger and thirst, or sound and sight, or the passage of time and the relations in space are experienced in the unique ways that they are. The neurobiology of consciousness is not a single problem in the way that the structure of oxytocin receptors, for example, is a single problem. It is a many-factored problem.

In a convergent way, Petar Dimkov (2020) mentions a great opposition between Eastern and Western theories of mind. The first, in general, postulate the non-existence of a *self*, while the second present countless variations, many of them assuming the existence of an entity that could be named as *I*. According to him, when discussing the subject, Eccles and Popper (1985, p. 108) would support the last perspective, as they point out, on the one hand, that "our personalities, our selves – are anchored in all the three worlds, and especially in World 3." On the other hand, authors such as Albert Newen, Leon de Bruin, and Shaun Gallagher (2018) bear that computational and materialist theses were predominant in the history of the philosophy of mind, particularly because they relegate a large part of the causal or ontological role of mental processes in the brain. Therefore, the *brainocentric* theses seek to justify that between the *self* and the *brain*, there is an extremely close connection, and there are important facts that can explain this mechanical relationship.

In this sense, we argue that, contrary to previous positions, the thesis of the extended and embodied mind represents an important step in refuting physical localizationism within the scope of brain processes (PERUZZO JÚNIOR, 2014, 2021). Thus, the mind is not just the pilot, as Plato put it, but it is an extension of various processes that extend intimately with the

environment. Therefore, this position indicates that the cognitive turn of this argument does not circumscribe mental processes to what occurs in the *bodily self*, as brainocentered arguments do not claim any identity that could make possible the emergence of self-awareness.

1 The self imprisoned in the skull

The physicalist conception that the brain structure, particularly its plasticity, can explain why an active psychophysical self is the programmer and coordinator of the brain is not sufficient to dissolve the consistency of the identity problem, which is the possible parallelism between physical and mental states. Jaegwon Kim (2010, p. 11), in *Philosophy of Mind*, states that "materialism, or physicalism, broadly understood is the basic framework in which contemporary philosophy of mind has been debated". In other words, according to Kim, the conception that all explanations about objects and phenomena in the world would be reducible to physical explanations continues to support several theories of mind, such as John C. C. Smart's Theory of Mind-Brain Identity (1959). It follows, therefore, that the physical properties capable of explaining mental processes would be the brain properties, which, then, confine the mind to the limits of the skull (KIM, 2010). However, this argument does not clarify how we should deal with the apparent unity between mental processes and conscious experience, as discussed above.

In this sense, some research on the theme of emotions, such as those carried out by Luiz Pessoa (2017), Dean Burnett (2018), and Ralph Adolphs and David Anderson (2018), insisted on the central role played by the brain, which Giovanna Colombetti and Eder Zavala (2019) called *affective brainocentrism*. Adolphs and Anderson (2018, p. 308), for example, argue that "[...] emotions are fundamentally biological phenomena [...] implemented by neural mechanisms that we can discover and manipulate with neuroscience methods"; therefore, a "science of emotion" does not need theories about subjective aspects and consciousness. According to Damásio (2012, p. 135), in *Descartes' Error*, the essence of an emotion is the "[...] collection of changes in the state of the body that are induced in a multitude of organs through nerve endings under the control of a dedicated brain system, which responds

to the content of thoughts related to a particular entity or event." Furthermore, Damásio (2012, p. 202) also states that the *I* – for lack of a more adequate word – "[...] is based on the activities taking place throughout the body, that is, in the body itself and in the brain."

In any case, even arguments, that assume a narrative identity of the self, have been colonized by localizationist and brainocentrist tendencies. Therefore, we need to separate hypotheses that explain the functioning of the cognitive system from those that are defined under the influence of exogenous components to extend such functions of the self. In relation to the "self" and the notion of personal identity, it is also possible to identify proposals whose focus lies on the biological apparatus of the brain, for example, David DeGrazia's work (2005, p.73) in the book *Human Identity and Bioethics*: "[...] we have defended the biological view, which holds that we are essentially human animals and that human identity consists of the sameness of biological life." Similarly, Galen Strawson (2013) states that both experiences can be applied as *brain activity* and what the author calls *minimal subject*, that is, the subject of the experience. However, is it possible that a physical structure, like the brain, emerges as something such an apparent unity?

In Churchland's conception, this time in *Brain-Wise:* studies in Neurophilosophy, published in 2002, the self simply emerges from the brain. More specifically, it depends on the following three factors:

> Such unity and coherence as there is in my conception of myself as a self depends on, among other things, these neurobiological facts: (1) my body is equipped with one brain, (2) body and brain are in close communication, and (3) activity in diverse parts of the brain is coordinated at a range of time scales, from milliseconds to hours. (CHURCHLAND, 2002, p. 61).

According to Churchland (2002), what most defines the self par excellence are the self-representation capabilities with which the organism is endowed. Representations, in this sense, would be nothing more than patterns of brain activity whose function is to contain certain information. This ability results in the formation of a neural map. Unlike the geographic map that we consult, however, Churchland (2013) clarifies that there is no *I* that undertakes to read this map. Her reductionist thesis is that the brain is the self itself, making mental life dependent exclusively on neural events.

Sandro Nannini (2018) argued, in a complementary way, that both intentionality (i.e., the experience endowed with content) and self-awareness and awareness of living the experience have neural correlates, as it is possible to verify the specific deficiency of these characteristics in individuals with injuries or syndromes. People whose cerebral hemispheres have been separated often identify whether a bar is positioned vertically or horizontally while claiming that they cannot see the bar. That is, they do not experience consciously (WEISKRANTZ, 2009).

John Barresi and Raymond Martin (2013), likewise, state that most philosophers of mind are currently materialists and that, therefore, the question they seek to answer about the *self* would refer only to the connections of social and experiential factors from the material basis of the mind, that is, the brain. According to them, "[...] in addition, the issue of self-unity is less of a problem if the organic dimension is taken as primary than if the experiential or social dimensions are taken as primary." (BARRESI; MARTINS, 2013, p. 33). In other words, for materialist and brainocentered theses, the question about the *self* would not be pertinent.

Thus, it seems evident that a number of studies from neuroscience seem to be in the search for neural correlates to explain the *self* and, consequently, its functions (such as the *first-person perspective*, self-image recognition or recognition of personal characteristics, etc.). The research by Kai Vogeley et al. (2004) and George Northoff et al. (2006) follows the argument mentioned above. Vogeley and Gallagher (2013) argue, however, that such investigations demonstrate that the *self* is simultaneously in all and none of the regions of the brain, as different correlates demonstrate activations in different brain areas, while these same areas present activations for activities not related to the *self*.

In any case, the brainocentric position has argued that the *self* is not a kind of homunculus that occupies a special place in the human brain. On the contrary, they insist on an approach that considers psychic life absolutely dependent and coincident with its physical constitution. This hypothesis, however, cannot deal with what would seem to give unity to a conscious experience. Thus, in the wake of this impasse, the perspective of the embodied mind locates cognition not as something that occurs in us in isolation but as

distributed and/or extended throughout this world. Cognition is, thus, embodied because it emerges from the body as a whole, going beyond the brain and placing itself in the shaping environment from which it emerges. Therefore, unlike physicalist and dualist proposals, embodied cognition dissolves the mythological timbre that inflates the concept of *personal identity* and, consequently, offers an alternative to the problem of consciousness by relocating cognitive processes as parts of the world.

2 From the enclosed self to the extended self in the world

Andy Clark's thesis (1993, 1997, 2003, 2008, 2016), known as the *extended mind thesis*, in contrast to the brainocentric conceptions mentioned above, postulates that cognitive processes are far beyond the skull and, therefore, the brain. According to him, human development could not occur in a brain isolated from an environment: "Mind is a leaky organ, forever escaping its "natural" confines and shamelessly mingling with body and with world." (CLARK, 1997, p. 53). The mind, the body, and the world would be deeply connected and mutually constituted constantly and inseparably.

This position is exemplified by the analysis of simple tools that we use in our daily lives, such as a pencil or a spoon. Some people may report formulating their ideas in a more elaborate and complex way when they are in front of a sheet of paper and a pencil, or a blank screen and a keyboard, than when they have no artifacts to act in the world. Hence, Clark (1993, 1997, 2003) calls, on the one hand, the tools we incorporate into our cognitive processes as *cognitive artifacts* or *cognitive extensions* and, on the other hand, the mind as *scaffolded mind* or *extended mind*.

However, the extended character of the mind follows an even more critical trail. Clark (1997, p. 69) argues, for example, that "[...] it would be comforting to suppose that this more integrated image of mind and world poses no threat to any of our familiar ideas about mind, cognition, and self. Comforting but false." That is, when we refer to the extended mind, we should also be conceiving an extended self. Therefore, the old notion of the organizing center of the mind would be illusory. Through ordinary impressions, we would be used to confuse the stream of conscious thought with what we commonly name as "I."

This stream of conscious thought is responsible for making broader decisions regarding the organism's system as a whole. However, it would not, in turn, be a separate entity, as we usually think. In other words, this thought would be able to formulate a narrative that we may call the *narrative self*, responsible for drawing the illusion that the *self* should be a special organization center (CLARK, 1993, 1997, 2003). In the author's words:

> My sense of myself as the protagonist in my own ongoing story is conditioned by my understanding of my own capacities and potentials—an understanding that must surely be impacted, in deep and abiding ways, by the technological cocoons in which my projects are conceived, incubated, and matured. (CLARK, 2003, p. 142).

The *self* would, therefore, be a *soft self* (CLARK, 2003, p. 134) and/or a *hybrid self* (CLARK, 2003, p. 56) immersed in a system of *hybrid cognitive circuits* (CLARK, 2008, p. 68). After all, it would encompass the supposedly external processes; moreover, obviously, it would always be blending in with the environment and its artifacts. Damásio, for example, in *Self comes to mind: constructing the conscious brain*, states that consciousness is produced by an *autobiographical self* – which is an autobiography that has become conscious – and by a *central self* – which pulsates incessantly and is always online. According to him, the autobiographical self "[...] matures, thanks to the gradual sedimentation and the re-elaboration of our memory." (DAMÁSIO, 2011, p. 260). Damásio's view, however, bears that to build the autobiographical self, the brain needs two combined mechanisms: First, to subside to the mechanism of the central self [which guarantees memories the possibility of being treated as objects and being conscious]. Second, a coordinated operation of the brain that involves steps such as relationships between memory and images, interactions of these ordered images with another part of the brain, that is, with the *protoself*, and, coherent results during a certain period. However, he rejects coordinating mechanisms as the center of consciousness or interpreting homunculi, as the results of every operation materialize "[...] in the brain structures that generate images and generate the mind." (DAMÁSIO, 2011, p. 263).

Gallagher (2014), from a perspective that also distributed and extended regarding the mind, as mentioned, argues that several dimensions need

consideration to explain the human functioning of the self: affections, motivations, sense of agency (i.e., who makes certain decisions and acts in the world is me), and sense of ownership (i.e., that a certain part of my body is mine), among others. Therefore, in line with Clark's defense that the mind is a system also constituted by the body and the world, the notion of intercorporeality needs to attack positions that enclose the self and cognitive processes in the cranial structure, as in Damásio's case. Thus, in other words, the embodied perception of the movements and the actions of others and oneself actively act in the construction of the *self*. This purports, therefore, to dissolve the ontological concept of identity as a property.

Thus, the extended mind thesis seems to propose that the notion of a system, although this generates the impression that it is an individuated system, is characterized by *ecological control* in the cognitive sphere. According to Clark, in *Soft selves and ecological control*, "[...] these larger problem-solving wholes, I would like to argue, are not simply extended cocoons for the 'real' selves, choosing agents and cognitive engines hidden deep within. Rather (or so I wish to suggest) they then are those selves, agents and cognitive engines." (CLARK, 2007, p. 104). The fact is that if the "I" is not a cocoon, our cognitive mechanisms are shamelessly diluted in the web of relationships that make up the world, as suggested by the phenomenological or functionalist reading of the extended mind. Therefore, an epistemic action should be understood as an action that changes the nature of cognitive tasks, because we make use of structures available in the environment (maps, language, culture, etc.) to perform certain tasks.

3 The fluid and dynamic nature of cognitive processes

If computational, physicalist, and/or materialist paradigms of the mind, figured as an expression of a brain-centrist reductionism (SEARLE, 1992; MALAFOURIS, 2013; CHAMBLISS, 2018; NEWEN; DE BRUIN; GALLAGHER, 2018), the advent of the extended and embodied theories of mind allows to radically redefine the way in which cognitive processes should be thought: *there is no demarcated border region between them and the environment*. In fact, the environment plays an extremely active role in cognitive processes. Obviously, this is not a relationship of dependence, but the constitution, as the organism and the environment purported here dissolve

the classic argument between the internal and the external. As Somogy Varga (2019) argues, it would be a *nature dependence* given that the character and nature of cognition depend on the body and the world. Therefore, this position represents a cognitive shift in relation to prospects that place the mind as a phenomenon attached to the skull or brain. However, this is only a part of the issue that needs to be addressed more rigorously. Clark (1989, 1993, 1997), for example, tries to oppose his thesis to Fodor's Computational Theory of Syntactic Image, seeking to reverse any conceptual distinctions between perception and action, mind and world, direct and indirect access, and so on.

For this reason, in the theories mentioned above (SEARLE, 1992; NEWEN; DE BRUIN; GALLAGHER, 2018), the brain is located as the locus of mental processes and representations and, in a broader sense, as a vehicle for mental content. However, if cognition can be extended and if it is in a process together with the environment, would mental states still be determined by brain ones? A physicalist answer, for example, would assert that representations are mental processes that, in turn, would have an identity relationship with brain processes. Clark, however, makes a mistake in arguing that the coupling of external objects to cognition is sufficient to explain their transfer of content. Thus, indirectly, the functionalist argument behind this position cannot indicate what the cognitive brand would consist of. Kenneth Silver (2021), for example, points out that members of a group could fulfill their role in the action without mentioning their mental states, as the behavior would not be caused by such states and the agent could simply respond to aspects of beliefs, desires, and intentions that are necessary for rational agency.

Now, if, on the one hand, the extended mind thesis seems to account for how artifacts contribute to expand and improve cognition, on the other hand, the argument cannot dissolve the criticism that could be attributed to the interaction and division between bodies, biological and inorganic artifacts. When Clark (2016) defends, for example, the notion of cycles between predictions and perceptions, it would be possible to argue that the boundaries between the interior and the exterior are still present, albeit in a *microfunctionalist* way, or, perhaps, according to Peruzzo Júnior and Stroparo (2020), in a *neofuncionalist metaphor*. However, the paradigm of embodied cognition, as we have argued, by positioning itself in a favorable way for the

hybridism between organic and cybernetic, also fails to show whether such tools can repair or replace damaged organs by assuming their cognitive functions. According to Malafouris (2013, p. 59), the success of embodied cognition allows for a new basis for the mind-body issue, as it recognizes that bodily characteristics play a significant role in the way the organism thinks and relates to the world.

In any case, the theories of the embodied mind share the defense of a rescue of the body and the environment because the psychological processes would be incomplete without the former. Visual perception, for example, would be a meaningless process if we did not consider the characteristics of the human body or even the stimuli that reach the specialized areas of the brain in the visual cortex, which is responsible for synthesizing the information received in terms of color, shape, texture, relief, and so on. In this sense, visual perception requires a body, just as a body is intrinsically linked to various functions. This could explain, for example, the extent to which direct experience about one's own mind and the minds of others would not be an internal attribution but a recognition that occurs from the emphasis on the stimuli shared between the subjects.

Another example, also associated with vision, is the *binocular rivalry* experiment reported by Jean-Luc Schwartz et al. (2012) and mentioned by Clark (2016). It is a *rivalry* in that a human subject is exposed to two images, each one arranged only for the perception of one eye. From this, if the image is, for example, a house and a face, the perception will not form a junction between them. On the contrary, the result will be an endless oscillation between the perceptions of one image and the other. The extended mind thesis would explain this through the aforementioned predictive processing cycles: predictions and perceptions take turns and interact continuously, one responding to the other, without being able to delineate a beginning or an end. Thus, the two images would take turns given that both do not cancel the forecasts made by the system.

In this sense, Varga (2019), in *Scaffolded Minds*: integration and disintegration, illustrates the role of the body and spatial conceptions appealing to the way we refer to people who are more receptive or more *emotionally distant* through the temperature, such as *warm people* and *cold people*, or to some uses of the notion of *heavy* to qualify issues, for example, a

heavy responsibility, the *severity of the situation*, and so on. In other words, we should accept that cognitive niche construction includes unloading, modifying, and incorporating environmental structures because they enhance cognitive abilities.

Thus, as mental phenomena could never be dissociated from such artifacts and tools, including language (LUPYAN *et al.*, 2020), the notion *all in the head* becomes a mistaken metaphor for positing that cognitive processes, in general, should be reduced to the brain. Therefore, it is not necessary to have mental representations to postulate the existence of the given cognitive content. These are possible simply because the idea of external reality is not the opposite of the symbols that we manipulate daily or the neurophysiological makeup of the brain. On that account, the intersubjective dimension of this argument ends up inhibiting the need for a homunculus to organize actions and behaviors, as the consequence is a *self* composed of this hybridism between perceptions, actions, and the world. What would, then, be left in place of the *self*?

In Andy Clark's conception of the extended mind (1997, 2003, 2008), for example, mental processes extend through objects and devices present in the world when they respect what the author, together with David Chalmers in the paper "The extended mind" (1998), names *parity principle*:

> All one needs is the very weak functionalism captured in the Parity Principle: roughly, if a state plays the same causal role in the cognitive network as a mental state, then there is a presumption of mentality, one that can only be defeated by displaying a relevant difference between the two (and not merely the brute difference between inner and outer). (CHALMERS, 2008, p. xv).

The consequence of this, as Clark (2008, p. xxviii) argues, is accentuated in the following statement: "In building our physical and social worlds, we build (or rather, we massively reconfigure) our minds and our capacities of thought and reason." Thus, from the moment we use a certain tool, in accordance with the parity principle, a *temporary whole new agent-world circuit* is formed (CLARK, 2008, p. 31). A fundamental aspect of the cognitive turn is, therefore, both the junction between action and perception and the partial and hybrid nature of mental representations. The understanding of numbers,

for example, simultaneously derives from words, that is, from language, and from visuo-spatial competences, more specifically, notions of quantity and relative location in a number sequence. In other words, finally "[...] what we perceive is determined not only by the physical properties of the current input, but also by our perceptual history." (LUPYAN *et al.*, 2015).

Therefore, Clark's functionalist externalism differs from traditional functionalism precisely because it adds external elements to cognitive functions. In other words, while traditional functionalism - the one that conceives a mental state as a causal function between sensory inputs and behavioral outputs, for example - challenged those thesis that reduced cognitive processes to neural states, the extended functionalism derived from Clark's thesis , in turn, expands this initial functionalist proposal. According to Wheeler (2010), Clark's thesis extends the cognitive status to external elements, since, according to the aforementioned parity principle, such elements have roles that are functionally equivalent to internal elements. Such functionalism, however, not only challenged the traditional, but also added to "classical externalism" - for example, the one proposed by Putnam (1976) - the active character of external elements. That is, while the *traditional* conceived intentional states as "relational properties", since they would always be referring to something external, the *active externalism* in Clark and Chalmers (1998) conceives certain objects and processes as having an active role in cognitive processes.

The implications of such a turn look promising. Varga (2019), for example, argues that the perspectives of the embodied mind offer psychotherapeutic studies and interventions that classical materialistic and/or computational proposals could not reach. Depression, from this cognitive turn prospect, can finally be contemplated in its entirety: states that somehow afflict the emotions, the body, and the phenomenal consciousness, which would show that it is not just about brain dysfunctions.

> This is where approaches that draw on EC [embodied cognition] could offer significant contributions to understanding mental disorders, analyzing them not merely as "brain dysfunctions" but as disturbances of an immersed embodied interaction with the environment, mediated by the brain. (VARGA, 2019, p. 9).

The cognitive turn carried out by the extended theory of mind, at this point, becomes self-explanatory: mental processes are not phenomena restricted to the brain and disconnected from the body and the individual's action in its environment. On the contrary, such processes explore different ways of performing from a certain functional profile that endorses, on the one hand, external components and, on the other hand, their functional extensions. The world has become a reservoir of relevant information, such as perceiving, reasoning, and remembering, among others, through which cognitive processes become hybrids; that is, they have a dynamic nature dependence and are not disconnected from the world, the action, and the context. Hence, it is not a radical dissolution of the self, but a recognition that human minds are in deep and important contact with the wider world, in the same manner as feelings, learning, and thinking are structured by the interactions of our body with the world around us. The *self*, then, assumes the form of a higher cognition, which is built on a substrate of embodied perceptual-motor capacities, since the action itself becomes simply the structuring of information and the ability to choose, nothing more than the internal architecture executed by the predictions.

Some might argue that this view of a dissolved *self* would automatically be refuted in the face of arguments that appeal to the notion of *qualia*, such as Nagel's (1974) classic argument in "What is it like to be a bat?" or that of Jackson (2012) in Mary and the perception of the redness. But there are also alternatives in the extended mind functionalism paradigm, whose dissolution takes place, we might argue, on the assumptions of predictive processing. In this sense, Clark (2019, p. 659) suggests that we should specifically ask "under what conditions would creatures like that infer that they are home to puzzling 'qualia'?" and adds that "we must here read 'infer' in an experientially neutral way". The inference of *qualia* would be the result of a predictive system that captures a series of details and whose complexity fails to emerge completely for conscious thought, which, for beings that express themselves in verbal language, for example, makes it difficult to explain what we perceive and how it is to have the experience of this perception. The discrepancy between what iscaptured and what emerges to consciousness would explain our "puzzlement" with *qualia*.

The question of *qualia*, for Clark (2019, p. 661), does not end at this point, however. As might be expected from a functionalist thesis, *qualia* would also play a role in the cognitive system, namely: "[...] capturing useful coarse patterns in our own behavior: ones that can inform our own reasoning and planning, and help us co-ordinate with other similarly designed agents." In the face of such arguments, selves come to be understood as a wider extended cognitive system, or even *hybrid* (MILOJEVIC, 2018), whose dimensions of both qualitative experiences and personal identity, for example, would not be threatened, since the sense of stability and continuity of the agent in question would remain preserved (GRIGC, NOVINA, 2022).

Conclusion

The main criticism of brainocentric perspectives refers, roughly, to the *explanatory gap* or *the hard problem of consciousness*, a term coined by Chalmers (1996; NANNINI, 2018). It is, as we have seen, the absence of an effective explanatory potential connection between neuronal activities and phenomenal experiences. Therefore, embodied cognition claims that the formation of representations does not work as an internal painting of the external world, as cognitive processes tend to follow simpler and more accessible solutions to show how we act in and with the world. According to this view, we are not facing two distinct domains (physical and mental) because we are, in the final analysis, *our bodies and the environment that we build and that act on us*. Therefore, *selves* are distributed agents, hybrid problem-solving sets, being essentially opportunists and explorers, transforming everything around them into fluent problem-solving routines and promoting guided activities and flexible adaptive responses.

In any case, some authors, such as Chalmers (1996) and Ned Block and Robert Stalnaker (1999), argue that this explanatory gap can never be overcome, as the reduction of phenomenal experience to physical or cerebral processes would be impossible. The *hard problem*, thus, would be how physical processes give rise to subjective experience, particularly as the concept of information would be ubiquitous and could transcend physical processes. Therefore, the theory of embodied cognition does not assume that cognitive processes have a purely physical domain, which also does not mean accepting that conscious experience and self occur "shamelessly" anywhere in the world.

How is it possible, we may ask, to conceive the "self" from the cognitive turn? The mind is a system designed to solve problems and optimize and amplify its action; thus, using available artifacts as effectively as possible and, therefore, incorporating them into a single mind-world circuit, "[…] minds like ours are transformed by the web of material symbols and epistemic artifacts." (CLARK, 2008, p. 57). The mind-body-world system becomes a *cognitive niche*, taking the brain from the prominence of the previous theses and putting it as a component. Human intelligence, therefore, develops through immersion in this niche along with the neural plasticity with which the brain is endowed.

PERUZZO JÚNIOR, L.; STROPARO, A. L. Dissolvendo o self: a virada cognitiva da teoria da mente estendida. *Trans/form/ação*, Marília, v. 46, n. 2, p. 193-214, Abr./Jun., 2023.

Resumo: O objetivo deste artigo é demonstrar como a teoria da mente estendida, particularmente os argumentos de Andy Clark, pode explicar os processos mentais não como fenômenos restritivos ao cérebro e endossar sua conexão com o corpo e o ambiente. Dessa forma, inicialmente, reconstroem-se as principais perspectivas materialistas que limitaram o *self* ao crânio; em seguida, aponta-se como o caráter estendido da mente escapa aos seus limites naturais e se mistura "descaradamente" ao mundo. Argumenta-se que artefatos externos desempenham um papel importante na orientação de ações, de modo que mudanças no ambiente podem causar mudanças no comportamento do agente cognitivo, configurando uma dependência constitutiva. Desse modo, a tese da mente estendida desafia tanto o funcionalismo tradicional quanto o externalismo, pois, por um lado, considera os processos cognitivos e os estados mentais como interações relevantes do indivíduo com o ambiente e, por outro, como comportamentos orientados pela intenção. Por meio da integração dos corpos biológicos com artefatos ou ferramentas, sustenta-se uma leitura que dissolve a clássica "lacuna explicativa" das ciências cognitivas.

Palavras-chave: Cognição. Self. Mente Estendida. Cerebrocentrismo. Andy Clark.

References

ADOLPHS, R.; ANDERSON, D. J. **The neuroscience of emotion**: a new synthesis. New Jersey: Princeton University Press, 2018.

BARRESI, J.; MARTIN, R. History as prologue: western theories of the self. *In*: GALLAGHER, S. (Ed.). **The Oxford Handbook of the Self**. Oxford: Oxford University Press, 2013. p. 33-56.

BURNETT, D. **The happy brain**: the science of where happiness comes from, and why. New York: Harper Collins, 2018.

CHALMERS, D. J. **The conscious mind**: in search of a fundamental theory. Oxford: Oxford University Press, 1996.

CHALMERS, D. J. Foreword. *In*: CLARK, A. **Supersizing the mind**: embodiment, action, and cognitive extension. Oxford: Oxford University Press, 2008. p. ix–xvi.

CHAMBLISS, B. The mind-body problem. **Wiley Interdisciplinary Reviews:** Cognitive Science, v. 9, n. 4, p. 1-14, 2018.

CHURCHLAND, P. S. **Brain-Wise**: studies in neurophilosophy. Cambridge: The MIT Press, 2002.

CHURCHLAND, P. S. **Touching a nerve**: the self as a brain. New York: W. W. Norton & Company, 2013.

CLARK, A. **Being there**: putting brain, body and world together again. 2. ed. Cambridge: The MIT Press, 1997.

CLARK, A. **Natural-born cyborgs**: minds, technologies, and the future of human intelligence. Oxford: Oxford University Press, 2003.

CLARK, A. Soft selves and ecological control. *In*: ROSS, D. *et al.* **Distributed Cognition and the Will**: individual volition and social context. Cambridge: The MIT Press, 2007. p. 101-122.

CLARK, A. **Supersizing the mind**: embodiment, action, and cognitive extension. Oxford: Oxford University Press, 2008.

CLARK, A. **Surfing uncertainty**: prediction, action, and the embodied mind. Oxford: Oxford University Press, 2016.

CLARK, A. Consciousness as generative entanglement. **Journal of Philosophy**, v. 116, n. 12, p. 645-662, 2019.

CLARK, A.; CHALMERS, D. The extended mind. **Analysis**, v. 58, n. 1, p. 7-19, 1998.

COLOMBETTI, G.; ZAVALA, E. Are emotional states based in the brain? A critique of affective brainocentrism from a physiological perspective. **Biology & Philosophy**, v. 34, n. 45, p. 1–20, 2019.

DAMÁSIO, A. **E o Cérebro criou o homem**. São Paulo: Companhia das Letras, 2011.

DAMÁSIO, A. **O Erro de Descartes**. Emoção, Razão e o Cérebro Humano. São Paulo: Companhia das Letras, 2012.

DEGRAZIA, D. **Human Identity and Bioethics**. Cambridge: Cambridge University Press, 2005.

DIMKOV, P. R. The Concept of Self in Eastern and Western Philosophy. *In*: INTERNATIONAL E-CONFERENCE ON STUDIES IN HUMANITIES AND SOCIAL SCIENCES, 5., 2020, Belgrade. **Conference Proceedings**. Belgrade: Center for Open Access in Science, p. 197–204.

DI PAOLO, E.; THOMPSON, E. The enactive approach. *In*: SHAPIRO, L. **The Routledge Handbook of Embodied Cognition**. Abingdon: Routledge, 2014. p. 68-78.

ECCLES, J.; POPPER, K. **The Self and its Brain**. Berlin: Springer, 1985.

GALLAGHER, S. The socially extended mind. **Cognitive Systems Research**, v. 25-26, p. 4-12, 2013.

GALLAGHER, S. Phenomenology and embodied cognition. *In*: SHAPIRO, L. **The Routledge Handbook of Embodied Cognition**. Abingdon: Routledge, 2014. p. 9–18.

GRGIC, A.; NOVINA, M. Extended self and identity over time. **Disputatio Philosophica**, v. 23, n. 1, p. 65-76, 2022.

HEERSMINK, R. Varieties of the extended self. **Consciousness and Cognition**, v. 85, n. 103001, p. 1-12, 2020.

HUTCHINS, E. **Cognition in the wild**. Cambridge: The MIT Press, 1996.

JACKSON, F. Epiphenomenal Qualia. **The Philosophical Quarterly**, v. 32, n. 127, p. 127-136, 1982.

JACKSON, F. XII – Leibniz's Law and the philosophy of mind. *In*: MEETING OF THE ARISTOTELIAN SOCIETY, v. 112, parte 3, 2012. **Proceedings**… London: The Aristotelian Society, 2012, p. 269-283.

KIM, J. **Philosophy of Mind**. 3. ed. Boulder: Westview, 2010.

LUPYAN, G. *et al.* Effects of language on visual perception. **Trends in Cognitive Sciences**, v. 20, n. 20, p. 1-15, 2020.

MALAFOURIS, L. **How things shape the mind**: a theory of material engagement. Cambridge: The MIT Press, 2013.

MATURANA, H. R.; VARELA, F. J. **Autopoiesis and cognition**: the realization of the living. Dordrecht: D. Reidel, 1980.

MILOJEVIC, M. Extended mind, functionalism, and personal identity. **Synthese**, v. 197, n. 5, p. 2143-2170, 2018.

NAGEL, T. What is it like to be a bat? **The Philosophical Review**, v. 83, n. 4, p. 435-450, 1974.

NANNINI, S. The mind-body problem in the philosophy of mind and cognitive neuroscience: a physicalist naturalist solution. **Neurological Sciences**, v. 39, n. 1, p. 1-9, 2018.

NEWEN, A.; DE BRUIN, L.; GALLAGHER, S. 4E Cognition: historical roots, key concepts, and central issues. *In*: NEWEN, A.; DE BRUIN, L.; GALLAGHER, S. **The Oxford Handbook of 4E Cognition**. Oxford: Oxford University Press, 2018. p. 3-18.

NORTHOFF, G. *et al.* Self-referential processing in our brain – A meta-analysis of imaging studies on the self. **NeuroImage**, v. 31, n. 1, p. 440-457, 2006.

PERUZZO JÚNIOR, L. Intentionality, Conceptual Content, and Emotions. **Revista de Filosofia Aurora**, v. 31, n. 54, p.833-847, set./dez. 2019.

PERUZZO JÚNIOR, L. **As Múltiplas Faces da Realidade**: Percepção, Linguagem e Cognição. Curitiba: CRV, 2021.

PERUZZO JÚNIOR, L.; STROPARO, A. L. Processos cognitivos e mentes estendida: uma metáfora neofuncionalista? **Natureza Humana**, v. 22, n. 1, p. 34-49, 2020.

PESSOA, L. A network model of the emotional brain. **Trends in Cognitive Science**, v. 21, n. 5, p. 357-371, 2017.

PUTNAM, H. The meaning of meaning. **Minnesota Studies in the Philosophy of Science**, v. 7, p. 131-193, 1976.

SCHWARTZ, J. *et al*. Multistability in perception: binding sensory modalities, an overview. **Philosophical Translations of the Royal Society**, v. 367, n. 1591, p. 896-905, 2012.

SEARLE, J. R. **The Rediscovery of the Mind**. Cambridge: The MIT Press, 1992.

SILVER, K. Group Action Without Group Minds. **Philosophy and Phenomenology Research**, v. 102, n. 1, 2021.

SMART, J. J. C. Sensations and brain processes. **The Philosophical Review**, v. 68, n. 2, p. 141-156, 1959.

STRAWSON, G. The minimal subject. *In*: GALLAGHER, S. (Ed.). **The Oxford Handbook of the Self**. Oxford: Oxford University Press, 2013. p. 253-278.

VARGA, S. **Scaffolded Minds**: integration and disintegration. Cambridge: The MIT Press, 2019.

VOGELEY, K. *et al*. Neural correlates of first-person perspective as one constituent of human self-consciousness. **Journal of Cognitive Neuroscience**, v. 16, n. 5, p. 817-827, 2004.

VOGELEY, K.; GALLAGHER, S. Self in the brain. *In*: GALLAGHER, S. (Ed.). **The Oxford Handbook of the Self**. Oxford: Oxford University Press, 2013. p. 111-138.

WEISKRANTZ, L. **Blindsight**: a case study spanning 35 years and new developments. Oxford: Oxford University Press, 2009.

WHEELER, M. In defense of extended funcionalism. *In*: MENARY, R. (Ed.). **The Extended Mind**. Cambridge: The MIT Press, 2010.

Received: 16/02/2022
Accept: 17/07/2022