AN EVOLUTIONARY ARGUMENT FOR THE EXTENDED MIND HYPOTHESIS

Marcello Ienca

Swiss Federal Institute of Technology, ETH Zurich
Marcello.ienca@hest.ethz.ch
Orbis Idearum, Vol. 6, Issue 1 (2018), pp. 39-61.

ABSTRACT

This paper will present an argument in support of the Extended Mind Hypothesis (EMH). The argument consists of two sub-arguments, an evolutionary and a systematic one. The evolutionary sub-argument advances that there is good empirical and analytic reason to suppose that the mind is designed by natural selection to extend. The systematic states that EMH is more consistent with the current system of science and scientific taxonomy. In force of the connection between these sub-arguments, I will conclude that EMH is epistemically advantageous compared to purely internalist accounts of cognition. Furthermore, this essay situates EMH in a broader landscape of philosophical and scientific ideas and theories about the mind.

1. EMH AND ACTIVE EXTERNALISM

The Extended Mind Hypothesis (EMH) is the view according to which it is arbitrary to state that the mind is only contained within the boundaries of the nervous system. Rather, it may extend to objects and processes whose local position is outside the nervous system.[23] Advocates of the EMH refer indeed to a challenging notion of mentality that addresses the question as to the division point between the mind, the body and the environment by promoting the view of Active Externalism.[24]

---

[23] Traditional EMH definitions are used to bring up the division point of the brain in place of the nervous system. For example, Clark & Chalmers (1998) notoriously presented EMH through the slogan «cognitive processes ain't (all) in the head». However this is a very liberal description, that virtually all neuroscientists would agree with, for none of them would doubt extra-brainy tissues such as the cerebellum, the spinal chord and the optic and olfactory nerves to actively drive cognition. In my view, therefore, the challenging issue here is whether or not the mind may extend to components outside the nervous system and even outside the organism itself.

[24] For Active Externalism see Clark & Chalmers (1998), Sneddon (2002) and Dartnall (2005).

Active Externalism (AE) is the view that proposes that external factors may play an active role in driving cognitive processes. It is easy to note how this view and the EMH are strictly related. AE implies that some objects or processes in the external environment are integrated by the mind's information processing in such a way that actively drives the functioning of the mind itself. EMH moves a step further and says that, due to their active role in driving cognitive processes, such external objects or processes can be seen as *extensions* of the mind itself. For this reason, advocates of EMH claim that the mind may also integrate bodily and environmental components, and call the resulting integrated system Extended Mind (EM).[25]

An easy way to appreciate EMH is taking into account the traditional mind-computer functional analogy, but in an upgraded form. The classical version of the analogy says that the mind/nervous system relation resembles the relation between a personal computer hardware and a system software: the nervous system, on the one hand, is hardware, i.e. the physical part of the computer; the mind, on the other hand, is software, i.e. a set of programs and related data installed in the hardware in order to provide instructions for the hardware to accomplish tasks. The upgraded version of the analogy, which I call Upgraded Functional Analogy (UFA), basically accepts the classical analogy but stresses that human minds and nervous systems are respectively software and hardware of a very particular kind for they always have active interface. According to EMH, indeed, the nervous system is not like an ordinary computer, such as the one I am using to write this paper. Rather, it is like a special computer whose motherboard is constantly connected with many other computer and peripheral devices through integrated circuitry (such as chipsets, RAMs, BIOSs, ROMs, internal and external buses, peripheral storage devices etc.).[26]

In fact, in conformity with its commitment to AE, EMH refuses to view the nervous system as a closed circuitry. Rather, the nervous system is seen as a system constantly having active interfaces that keep it connected to

[25] Note that AE and EMH are not biunivocally related. Their relation is rather univocal: one might accept AE and nevertheless deny EMH; instead one might not consistently accept EMH and deny AE. This relational univocity will be further stressed later on.

[26] According to a classical definition, a hardware interface is described by the mechanical, electrical and logical signals at the interface and the protocol for sequencing them (Blaauw et. Al., 1997). Obviously the devices should not be always of the same kind, nor there should be a defined number of connected devices. Rather they might change in quality and quantity over time and local position of the computer, as the things that the mind interacts with also change in quality and quantity over time and local position of the brain. What counts for the sake of the analogy is that there is no time in which the computer is completely unplugged, as there is no time in the lifetime of an organism in which its brain is a closed circuitry and the mind does not interact at all.

other bodily components and the environment.[27]

Just as the nervous system has hardware interfaces, so the mind, in conformity with this view, must have software interfaces. Just as in a PC applications and programs running on the operating system may interact via streams or methods (and, in the online mode, also with the Internet), so the mind interacts both with other minds and with programs running on the external environment. Consider, for example, two people having an argument, or a student using a calculator to solve an arithmetical operation. In both cases the mind has an active software interface with external processes and interacts with them: in the former case we have a mind-mind interaction through the language program; in the second case the calculation program running in the student's mind interacts with an external calculation program installed on the calculating machine.

For EMH proponents, interactions of this kind characterize our overall cognitive life and can be found far and wide in our world. For instance, consider the case of a market trader writing on a slip of paper her customers' numbers, a child counting on her fingers, a girl using Google Maps on her iPhone to reach a certain place, a student writing down conceptual schemata to better succeed in the exam. In all these cases the mind establishes an interaction with some external objects (the slip of paper, the iPhone etc.) by running some programs that, through software interfaces, are capable to go across the mind and the external objects themselves. In order that such a process can run, external objects must get integrated within the same cognitive loop of the mind. This does not require only software interfaces; instead, also the physical interfaces of the nervous system and, to a certain extent, of the rest of the body, are likely to be used. The whole process resambles how a computer integrates external software and peripheral devices in its circuitry by means of both software and hardware interfaces.

On the grounds of this general view, advocates of EMH address the following question: given that the mind interacts with objects in the environment and integrates them within the mind's processing loop, do these external objects count as extensions of the mind?

In the light of the analogy presented above, the most plausible answer seems to be "yes, they do". In fact, it seems difficult to deny that if you connect your computer to an external device, the device does not become an extension of the computer. For example, if I connect a 3 TB hard disk drive to the 320 GB HDD laptop I am using to write this paper, I obtain an in-

[27] The easiest way to appreciate this is looking at the connection between the brain and the rest of the body: hundreds of thousands of complex neural pathways interconnect the brain with all other bodily regions, such as with the circulatory and the muscular system, and, through the sensory organs, with the external environment.

crease of approximately 10^3 in the original computer memory. Such a massively increased memory seems to be a property of the whole integrated system (computer + hard disk drive), and not of only one component.

Suppose I want to store somewhere the complete collection of Stanley Kubrick's movies and there is no sufficient space on my internal hard disk, so that I have no option but to use the external drive. In such a case, the data storage process can be executed only by the integrated system constituted both by the computer and the drive. The computer alone could not execute it, for it would not have sufficient memory space. And the drive would never be able to go without the computer in running the memory process.[28] In order to execute the process, computer and drive cannot be decoupled. This implies that such a process, namely the memory storage of Kubrick collection, is not a function of the sole computer, but rather of the entire computer-drive coupled system.[29] Therefore, it seems reasonable to argue that, in such kind of cases, the external drive becomes part of the computer, or, put it differently, that the computer extends to the drive. It is worth to point out that this is a definition computer scientists widely agree upon without big theoretical problems. Indeed in computer science, to say that plugged-in peripheral devices are parts of the computer, and interfaced softwares –while executing the function for which they are interfaced– are parts of the running program, are both ordinary expressions.[30] Furthermore, files containing programming that serves to expand the capabilities of or data available to a more basic program, are typically called software extensions. Taking UFA seriously, advocates of EMH argue that, insofar as the computer extends, there is no theoretical impediment to claim that the mind also extends to the rest of the body and the environment.

1.1. Historical roots

Seen through the lenses of the history of ideas, AE –and to some extent EMH too– have deep historical-philosophical roots. The idea that the external environment may play an active role in driving cognitive processes characterizes two central perspectives in 20th century philosophy: Merleau-

[28] As suggested in the book *101 Uses for a Dead Computer*, the best possible use for an unplugged hard drive is probably that as table tennis paddle.

[29] It is worth to point out that not every coupling can be considered an extension of the original processing unit. As Clark & Chalmers (1997, p. 6) put it, «for coupled systems to be relevant to the core of cognition, *reliable* coupling is required». How a reliable coupling should look like will be discussed in the next sections.

[30] See for example the entry "Extension" on Webopedia: <webopedia.com>.

Ponty's theory of perception and Edmund Husserl's phenomenology.

In his "Phenomenology of Perception", the French philosopher questions the existence of a clear-cut separation line between the body's interaction with the world and the cognitive dimension. Repudiating any purely cognitivist ontology of the mind, purely gnoseological description of the self and the late-Platonic idea of mind-body dualism, Merleau-Ponty described the "lived body" as the unity of what he called "a mind-body-world system". In his view, the "lived body" is better characterized as a complex system where the cognitive dimension of experience (including consciousness) are embedded in the body. In Merleau-Ponty's words: "My body is the pivot of the world: I know that objects have several facets because I could make a tour of inspection of them, and in that sense, I am conscious of the world through the medium of my body".[31] On these grounds, Merleau-Ponty hypothesized that body alterations may affect a person's phenomenal experience.[32]

Similarly, in Husserl's phenomenology, humans are seen to perceive, understand and interact with the world as "embodied subjects" (Husserl, 1950). The body as described by Husserl is neither internal to consciousness nor external to the subject in the environment, it is neither subject nor object, rather it is "a thing 'inserted' between the rest of the material world and the 'subjective' sphere" (Husserl, 1950). Through these lenses, the body becomes the instrument that orientates humans through their experience of the world and cannot be easily decoupled from the mind.

It is worth noting that, according to Husserl's phenomenology, the body is not understood as the objective and material body (in German, *Körper*) but rather as the subjective body (*Leib*). Such subjective body is seen as something that constantly and actively interacts with the external world (Walsh, 2017). Wherever there is a body, there is a personal interaction with the world. This notion of body as *Leib* can be dated back to Friedrich Nietzsche's *Also Sprach Zarathustra*, where it was used to highlight the poietic function of the body in shaping subjective experience. In a section called "The Despisers of the Body", Zarathustra argues: "Ego, sayest thou, and art proud of that word. But the greater thing—in which thou art unwilling to believe—is thy body with its big sagacity; it saith not "ego," but doeth it (Nietzsche, 1914).

This historically stratified philosophical stance was imported into the cognitive sciences through the contemporary philosophy of mind, as several authors have characterized the "lived body" as "the way the body structures our experience" (Gallagher & Zahavi, 2007). Phenomenological accounts of cognition have found support in two specific accounts of AE called, respec-

[31] Merleau-Ponty, M. (1945).

[32] Bullington (2013).

tively, *embodied cognition* and *environmental cognition*. According to the embodied framework, "cognition is deeply dependent upon features of the physical body of an agent" (Wilson & Foglia, 2011). Proponents of ecological cognition, in contrast, highlight the active role of environmental factors.

2. THREE OBJECTIONS AGAINST EMH

EMH was originally formulated at the end of the '90s by two philosophers, Andy Clark and David Chalmers, and obtained a certain success among philosophers sympathetic to an externalist account in theory of meaning.[33] In contrast, it was often challenged by philosophers of mind confident of the existence of an intrinsic content of mental states.[34] The uptake of EMH in the mind and brain science also appears quite limited. Indeed, the default brain-centric position in neuroscience is that there is a well-defined system responsible for cognition, to virtually the same extent to which there is a well-defined system, the cardiovascular, responsible for blood distribution. This system is the nervous system. Extra-neural systems might, under certain circumstances, play a causal role but they do not become constitutive of cognition or any other brain-enabled function.[35]

But why this hostility? If, in the language of computer science, peripheral devices connected to a computer do count as *extensions* of the computer, and programs interfaced with the computer do count as *extensions* of the computer operating program, why does not go the same for brains and minds?

The most instinctive objection is that simply showing that computers seem to extend to external components, still does not say much about whether also the mind extends. This objection has some appeal. The upgraded analogy, indeed, is just an analogy, and analogies must not always be taken seriously. For example, molecular biology handbooks often present the process of creating a complementary RNA copy of a sequence of DNA through the analogy with graphic code transcription. But such an analogy is only used for explanatory aims, without any commitment to the view that what is true for graphic code transcription must be necessarily true for DNA transcription too. And for sure no geneticist working on transcription would

---

[33] Externalism about meaning, also called semantic externalism, is the view that the meaning of a term is determined, in whole or in part, by factors external to the speaker. See Putnam (1975).

[34] See Adams & Aizawa (2001), Aizawa (2009).

[35] It is easy to note that the default position in neuroscience, though apparently in contradiction with the EMH, is manifestly compatible with AE.

consult a code compiler. But simply answering that way would just mean begging the question, for the salient problem here is precisely why UFA should not be taken seriously. Opponents of EMH, therefore, use to enforce this intuitive objection by advocating the three following arguments against UFA and the EMH:[36]

A. Computers and their extensions have, to a certain extent, a similar physical implementation; the nervous system and, say, an iPhone, are realized in two completely different ways.

B. Whereas a hard drive has a phisical connection point with the computer, environmental objects that interact with our mind, usually do not.

C. Whereas the computer and the hard drive are designed by engineers to be coupled in an integrated system, and the programs they run are programmed by computer scientists to have a reliable software interface, the connection between the mind and the objects in the external environment, on the contrary, is not designed but merely contingent.

2.1. Objection A and the problem of the physical implementation

Objection A has the following structure:

P1: In order for one system S2 to be considered as an extension of a system S1, S1 and S2 must share the same physical implementation.

P2: External objects and the mind do not share the same physical implementation.

C: Therefore, external objects cannot be considered as extension of the mind.

This argument has at least three strengths. First, it is logically well-formed. Second, it projects an induction (P1) resting upon good empirical evidences. Indeed, most things commonly referred to in our ordinary language as constitutive of something, i.e. as its extensions, seem to share the same physical implementation of this something. For example, the extension to a house usually has a similar physical implementation of the house itself: it could be made of bricks, or cement, but it would be very strange if it were made of optic rays. P1 infers thus something apparently true about how constitutive relations are. Third, A states something *prima facie* self-evident about the metaphysics of the extended mind cases. Indeed, at least at the actual state of affairs of biotechnology, there are no objects sharing the same physical implementation of our mind as artificial neural tissue is far from enabling human-like cognition.

[36] To simplify the index I will label these objections with the first three letters of the Latin alphabet.

However, from the perspective of EMH, simply pointing to the implementation medium as justification (as in P1) is of no use at all to explain the nature of a certain function. Rather, it begs the question. The core idea of EMH is precisely to call into question what the implementation medium of cognition really is, i.e. whether it is only the nervous system or the active coupling between the nervous system and something else outside of it. As Clark & Chalmers put it, «we must find some more basic underlying difference between the two» than the mere implementation medium.[37]

2.2. Objection B and the Parity Principle

In appealing to «some more basic underlying difference» the advocates of EMH typically refer to a functional difference. Indeed, one further theoretical feature of EMH, besides a commitment to AE, is the endorsement of functionalism. In its classical definition, functionalism is the doctrine that what makes something a mental state of a particular type does not depend on its internal constitution, but rather on the way it functions in the system of which it is part. In the case of EMH, it is the commitment to the view that what makes something a constituent of the mind does not depend on its internal constitution, but on whether this something *functions* as if it were a constituent of the mind within a given cognitive system. This special instance of functionalism is commonly known as parity principle (PP), and counts, in conjunction with AE, as the conceptual basis for EMH.

In the words of Clark & Chalmers, PP says: "If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, than that part of the world is part of the cognitive process" (Clark & Chalmers, 1997, 3).

This implies that if something outside the skull functions as part of an intra-cranial cognitive process, then it should be considered as part of that very cognitive process. Consequently, if the system executing the exogenous process functions as part of the system executing the endogenous cognitive process, then the exogenous system must be considered as an extension of the endogenous one. The only condition of equivalence that PP postulates is for S1 and S2 to execute the same, or at least, a highly similar function. In force of this fact the advocates of EMH are in a good position not only to counter the objection A but B too.

Let us see how: B rejects UFA on the grounds that whereas a hard drive has a physical connection point with the computer, environmental

[37] Clark & Chalmers (1997, 10).

objects that interact with our mind, usually do not. While A appealed to the implementation medium, B appeals both to the local position of S1 and S2, and to the type of coupling established between them. Thus according to B, in order for a system S2 to be considered as extension of a system S1, it must share the same local position of S1 (or, at least, be adjacent to it); and be interconnected with S2 in a way that resembles the interconnection between a computer and a peripheral device. Under these conditions it seems that a system S2 can count as a mind extension if and only if it get implanted somewhere in the central nervous system. Brain implants and neuroprosthetics seem to be the only systems that fulfill this condition. Therefore, why it can be said that neural implants are extensions of the mind, the same cannot be said for smartphones or other digital or non-digital technologies that do not share with our brain a specific point of physical connection.

However even this objection does not seem to dismiss the parity principle since non-functional properties like local position and type of interconnection are considered by proponents of EMH as irrelevant in determining whether a system S2 might count as extension of S1 or not.

2.3. *Objection C and the reliability of coupling*

Unlike A and B, this latter objections cannot be dismissed only by appealing to the basic conceptual assumptions of EMH. Rather, to counter C a good defense of EMH must also involve i) a collection of empirical data; ii) a general reflection on theory choice in science. However, unlike against A and B, the advocates of EMH did not present so far any significant counter objection against C.

Objection C rejects UFA and, consequently, EMH, on the grounds that, whereas the computer and the hard drive are designed by engineers to be coupled in an integrated hard system, and the programs they run are programmed by computer scientists to have a reliable software interface, the connection between the mind and the objects in the external environment, is not designed. Rather it is merely contingent. Therefore, C states something that cannot be dismissed simply by referring to PP. In fact, this objection postulates, in conformity with PP, that in order for S2 to count as extension of S1, S1 and S2 must be integrated in the execution of the same function, i.e. of the same cognitive process. This integration, however, cannot be of any kind, for not all mind- environment couplings count as EM cases. Imagine, for instance, to cast your mind back over your ex-girlfriend after seeing her car parked in a car park. Should then the car count as part of the recall process in your memory, and thus be considered part of your mind? For cases like these, the intuitive answer seems to be „not at all“. The reason why such cases are unlikely to count as EM cases, seems to be that in such cases the coupling between environmental items and the mind is unreliable. Precisely this reliability of couplings is what C points at issue: in order a process in S2 functions as part of a process in S1, there must be a reliable coupling between S1 and S2.[38]

But what does it mean for a coupling to be reliable? According to C in order a certain coupling to be reliable it must not be accidental or transient. Rather it must be designed for executing the function actually executed by the matching coupled system. Hardware and software interfaces are non-random in the strong sense that they are designed, built and programmed by technicians precisely to execute the function that they actually execute. They thus display a good functional reliability. Our ex-girlfriend's car, on the contrary, is designed to provide mobility but not to recall items or events. Therefore its coupling with the memory process is accidental and transient. But what about the EM cases mentioned in the first section? It seems not controversial to claim that, say, the iPhone's phone-book is designed to store memory data, in particular phone numbers, for it is built and programmed precisely to execute this function. Similarly an external device is designed to interface the PC in the execution of a given function. So the brain-iPhone interface seems to meet the reliability requirement.

One could argue that whereas the iPhone is designed for getting coupled with the brain, the brain is not designed for getting coupled with the iPhone. In other words, unlike the computer-HDD interface is bidirectionally designed, it seems that all genuine EM-cases are just unidirectional. So whereas many environmental objects, chiefly technological artifacts, are designed for getting coupled with our mind, it does not go the same the other way round. What sense does it make to say that our mind is designed for coupling with a technological artefact? This question sounds instinctively a bit odd, as notoriously no engineer designed our mind. Nevertheless, there is an important sense in which the answer to this question should be yes. Next section will suggest that our minds may have been designed for extending to objects and processes in the external environment to a similar extent to which computers are designed for extending to hard drives.

---

[38] Remember that the reliability of coupling, as noticed in the previous section, must be defined at the functional level. Hence, it can not depend on features such as similar physical implementation, local position and type of coupling

3. AN EVOLUTIONARY ARGUMENT FOR THE EMH

Natural selection is the immanent nonrandom force regulating biological evolution. It determines hereditary biological traits to become either more or less common in a population as a function of differential reproductive success of their bearers: in the course of generations, traits that became more common tend to have stable presence in a given genetic pool; whereas traits became less common tend to be eliminated from the gene pool and go extinct.

Although genetic mutations are random, the mechanism due to which mutated traits either spread in the gene pool or go extinct is typically nonrandom. Indeed it matches the following law: traits causing the organisms a better fitness for the environment where they live, tend to spread across generations; traits decreasing the fitness of an organism in that environment tend to go extinct. Traits of the first kind are called adaptations and their fitness coefficient is calculated through their reproductive success. For this reason, natural selection implies that virtually all hereditary biological traits in the actual biosphere are supposed to determine a certain amount of fitness advantage, or at least not to determine disadvantage; otherwise they had gone extinct. In operating this way, natural selection shapes organism phenotypes and, consequently, the matching gene pools. In conformity with the established lexicon of evolutionary biology, evolution designs organisms and their functions for making them adapted to a certain environment, as it, over long intervals of time, selects the genes responsible for organism phenotype and behavioral functionality.

My argument here is that the mind may have been selected for extending to the external environment. This would imply that the coupling between the mind and things in the environment is functionally reliable in the strong sense that it is non-random, but designed by evolution for executing adaptive functions, whose implementation increased our fitness coefficient in our environment.[39]

This argument moves in two steps. First, I will show why EM should evolve. Second, I will show why it is reasonable to call such a evolutionary trait EM.

[39] With this argument I am not necessarily committing myself to the claim that the property of being designed for equals the property of being reliable -and I am confident there might be cases that could falsify this claim. My argument indeed is also compatible with the more moderate claim that being designed for is sufficient for being reliable.

3.1. *The Natural Selection of Cognitive Faculties*

As previously pointed out, in order one trait to evolve through natural selection it must turn out to be adaptively advantageous. Being adaptively advantageous, however, is a variable criterion. One trait may indeed be advantageous in many respects. It may, for example, simply produce the minimal advantage of optimizing the organism's internal structure or function, such as better regulating its metabolism. Or it may determine a wide range of favorable adaptive outcomes (metabolic, ecological, behavioral etc.). The more complex a trait is, in more respects it must result adaptively advantageous. In recent years, evolutionary biologists have set the following list of adaptive criteria that a certain cognitive faculty should meet in order to make its selection evolutionary predictable:[40]

- Optimization in system internal organization
  - Optimization in system metabolic equilibrium
  - Optimization in input processes
  - Positive feedback on other system faculties
- Optimization in system output patterns
  - Refinement in output accuracy
  - Refinement in behavioral patterns
  - Increase in interaction success with the ecological niche

My claim that our mind may be designed for extending must thus be proven by showing whether or not EM meets the criteria listed above. This will be at issue in the next paragraphs.

3.1.1. *Internal organization*

The ability to extend some cognitive processes to the external environment may have determined an optimization in our internal system organization in three important ways. First, it may have produced a better metabolic equilibrium. Cognitive systems are dissipative systems that get pushed into operation by harnessing energy from a variety of metabolic pathways. The human brain, in particular, is one of the most dissipating systems of the biosphere, for it claims only 2% of our body mass, but is responsible for approximately 20% of our body oxygen consumption. For a cognitive system, therefore, energy must be constantly available for work (such as mechanical work) or for other processes (such as chemical synthesis and anabolic processes).

However, energy is for a system not always easily available. Food, for

[40] This list is a personal elaboration of Geary (2005) and Striedter (2005).

instance, our best resource to assimilate some of the essential nutrients that our cells convert in energy, is often scant. For this reason, evolution has favored those organisms capable to spark their life-maintaining processes with the lower possible expenditure of chemical energy. According to a basic principle in bio-energetics thus, all living systems try to execute their biological processes with the smallest effort/profit ratio, namely to obtain the best possible outcome with the lowest possible energy expense. Now, it seems that one of the easiest ways the nervous system has to achieve the smallest effort/profit ratio is by transferring some processes from the neurons to the environment, as the environment does not draw on our internal energy resources. In integrating some external objects in the cognitive loop, the mind off-loads onto these objects not only a certain amount of information, but a certain amount of chemical energy necessary for that information processing too. It thus behaves like a car exploiting a slope to reduce gasoline consumption. This evolutionary conjecture may accurately be proven experimentally. By means of in vivo magnetic resonance spectroscopy (MRS), indeed, we can follow the metabolic pathways of energy production (as glucose oxidation) and work (as monitored by the cycling of glutamate and GABA neurotransmitters). Furthermore, by using functional magnetic resonance imaging (fMRI), we could measure and localize small differences in neuronal activities between the extended and the unextended mode. Although specific neuroimaging experiments are still required, it seems reasonable to predict the cognitive off-load to things in the environment to determine a lower energy consumption for the same reason why driving a scooter for 10 miles determines a lower energy consumption than walking for the same length, as the process is mostly executed by the external system.

Behavioral biology and ethology also offer interesting data. Take for instance honey bee communication. Foragers honey bees use two sign language systems to communicate with their hive mates. By means of one of these languages, the so-called waggle dance, they share information about the direction and distance to patches of flowers yielding nectar and pollen, to water sources, or to new housing locations.[41]

Not all shared information is internally processed, for information such as the direction to the source is a function of the bee movement path and the hive spacial position. Instead of encoding complex (and metabolically expensive) information, forager bee's evolution has most likely favored the off-load of such information to physical objects in the environment and their subsequent functional re-integration in the cognitive loop. While performing the waggle dance, the bee's mind extends to the hive, for the

[41] von Frisch (1967).

hive becomes constitutive of the process executed by the bee's mind. This adaptive off-load is supposed by many ethologists to be very close to the smallest effort/profit ratio evolutionary attainable, for it enables the bees to accurately share fundamental information with the lowest possible level of neural complexity (hence, with the lowest possible expenditure of chemical energy).[42]

The same goes for some cognitive systems of mammals, such as the bat echolocation system. Bats emit calls out to the environment and listen to the echoes of those calls that return from various objects near them. They use these echoes to locate and identify the objects and themselves. This is also a clearly extended process, for it works if and only if bats exploit the external environment to reflect acoustic waves. Executing the same function internally would have been evolutionary too expensive, for it would have required the implementation of an internal surface acoustic wave reflector.[43]

Optimization of the internal organization does not operate exclusively at the biochemical level, but at the functional level too. In order for a system to be functionally optimized in evolutionary terms, hence to have a high statistical probability of propagating to next generations, it must turn out to be able i) to execute more functions than its unoptimized matching system; ii) to execute the functions of its unoptimized matching system more efficiently.

I would argue that in the case of EM both conditions are satisfied. In the first place, through extending to the external environment, the mind become capable of executing more cognitive functions than if it were confined within the boundaries of the skull. For example, when the human mind is coupled with an ultrasonography system, as it often happens with the minds of oceanographers, it becomes capable of processing a kind of information, i.e. ultrasounds, that it would not be able to process otherwise, for their frequency is greater than our mind's biological upper processing limit. More commonly, when we couple our mind with a computer connected to the internet, as we use to do every day, then we make available to our mind a huge amount of data –such as e.g. what the weather in Hanoi is like, how to get to the next subway station, how many calories a banana has, what Fellini's first film was etc.– that the decoupled mind would usually not have access to. This functional advantage does not only pertain to couplings with sophisticated technologies, but to simple artifacts and even to parts of the physical body too. All school

[42] von Frisch (1967), Seeley (1995)

[43] Teeling et Al. (2000)

students know, for instance, that thanks to the use of pen and paper people can perform complex arithmetical, algebraical and geometrical operations that they struggle more to solve if they would only lean on internal resources. The same goes for children counting fingers on their own hand[44] as well as for the cognitive features of many non-human species. As we have seen, honey bees would not be able to inform their mates about the source direction if they could not lean on external resources such as the hive and the sun position. Therefore, it is likely that evolution has favored those organisms whose mind was able to extend to the external environment in order to gain not only metabolic equilibrium, but cognitive improvement too.

Extended systems may also work more efficiently than their matching unextended systems. When it comes to cognitive processes 'efficiency' is basically calculated in terms of processing speed and outcome accuracy.[45]

Extended systems turn out to be more efficient than unextended ones, as they enable, under many circumstances, to process information faster and to produce more accurate outcomes. Kirsh & Maglio, for example, notoriously calculated that the physical rotation of a shape in the computer game Tetris goes about three times faster than the mental rotation of the same shape.[46]

This reveals external processing to be, at least under some conditions, dramatically faster than internal processing. The same can be said, again, for mathematical operations. Just try to do your maths homework both with and without calculator (or pen & paper set): you will suddenly notice a dramatic difference in the time it takes to work them out. Moreover, leaning on external supports does not only increase the processing speed, but the outcome accuracy too. To show this it is not difficult to find mental experiments taken from our every-day life: compare the accuracy in the process of dealing the correct phone number of your boyfriend both by trying to bring it to your mind and by finding it in the phone-book of your mobile phone.[47]

It is worth to point out that it is a common place in evolutionary biology to consider processing speed and outcome accuracy as fundamental features in determining the fitness rate of an organism. Suppose two different homo sapiens populations were both migrating from an arid to a

---

[44] See Ginsburg (1989), Dehaene (1999), Carpenter et al (1999).

[45] Cowan (1998), Piccinini & Scarantino (2010).

[46] See Kirsh & Maglio (1994). The exact ratio reported in the study was about 300 milliseconds of 1000 milliseconds to rotate the same shape through 90°.

[47] If you still think that you get the same accuracy you are invited to extend the experiment to all contacts saved on your phone-book.

prosperous region along the same track; they both know that at some point of the track there is a life- threatening danger like a colony of malaria-carrying anopheles or a pack of hyenas, but only one of the two groups used external cognitive strategies such as tracking the course on a hand-held support or marking trees and stones to remember the exact location of the danger. Given that leaning on external resources increases outcome accuracy, it is more likely for the group tracking the course or marking trees to survive and thus reach the prosperous land than for the group leaning exclusively on internal memory. Therefore, it is in principle likely that evolution has selected those individuals able to exploit external resources by integrating them within their own mind's cognitive loop.

One further way the natural selection of extended minds may have occurred is by sparking off a positive feedback effect on internal cognition. This means: developing some extended cognitive strategies may improve the effectiveness of the unextended processes too. A good example is natural language.[48]

Evolutionary psycholinguists agree that, since verbal communication first appeared about 120.000 years ago among the individuals of our species, it suddenly exerted a massive impact on our putatively non-verbal processes such as categorization, memory, visual discrimination, and even simply detecting the presence of a stimulus.[49]

From the point of view of EM this means, that a putatively extended cognitive process may not only be effective as such, but may have increased the effectiveness of internal processes too. It is, therefore, reasonable to argue that evolution might have favored those organisms that are capable to implement cognitive processes by exerting positive feedback on other processes. Another instance of this evolutionary phenomenon is the improvement in the accuracy of beliefs in consequence of long-term memory development. This phenomenon occurred gradually in the evolution of the human species during the Pleistocene, in consequence of the growth of hippocampal volumes, and can still be observed in Alzheimer's and Parkinson's disease patients. In these patients, memory reduction

[48] Language is a typical extended process, for it enables to off-load amounts of information from the internal resources of the mind to the outside, in particular to the social community. It is no coincidence, indeed, that the most complex communicative systems are found among eusocial species such as great apes and apodiploid insects (ants, wasps etc.). In addition, being language at least in part an adaptation, language evolution is supposed to meet the same principles for the evolution of cognitive faculties listed above. The seminal paper on language evolution is Pinker & Bloom (1990). See also Pinker (1994), Dunbar (1996), Deacon (1997).

[49] For a recent and updated study on this topic see Lupyan (2012). Also see Gallistel & Gelman (1992), Clark (1998), Carruthers (2002), Dascal (2002), Li & Gleitman (2002). For an exhaustive overview over this topic see Carruthers (2008).

caused by loss of neurons and synapses in the cerebral cortex systematically leads to a deficit in belief accuracy. For example, a patient prone to forget, say, the name of his daughter Amy, will also be more likely to have the false belief that her daughter's name is Laura. Insofar as it massively expanded our memory by off-loading amounts of information to external resources, the evolution of EM has also improved the accuracy of our beliefs. For example, the probability of having the true belief that Ashgabat is the capital of Turkmenistan is significantly higher when our mind is reliably coupled with external online resources such as Wikipedia, than in the decoupled mode. Now, it is not difficult to see how increasing the probability of having true beliefs is adaptively favorable.

Just consider the case of beliefs such as that cobras are venomous snakes, that tetanus infection occurs through wound contamination, or that inbreeding leads to a higher probability of congenital birth defects. Having such true beliefs as consequence of coupling our mind with external artifacts (before Wikipedia this function was executed by books, codes and norms) to attain shared social knowledge was for our ancestors a dramatic adaptive advantage, for it enabled them to avoid potential harms. Still nowadays, coupling our mind with external resources such as books, computers, websites etc. enables us to massively increase our survival rate, since it provides fundamental information for supporting life-maintaining processes, such as information about healthy diet, child care, how to avoid the transmission of infectious diseases and so on.

3.1.2. *External Organization*

In order to become a stable adaptation a cognitive faculty such as EM must not only guarantee an optimization in the system internal organization but in system output patterns too. In fact, what ultimately counts in evolutionary terms lies in behavior production, as evolution selects organisms for how they behave. This does not imply the system internal processes to be unimportant. Rather they are evolutionary relevant as they produce different behavioral patterns upon which natural selection acts by determining differential reproductive success. If a well-optimized cognitive process was not correlated to certain behavioral patterns, it would have no evolutionary significance. Vice versa no complex adaptive behavioral pattern may be independent of internal processing. Faster and more accurate input processing is, therefore, evolutionary significant for it determines faster and more accurate behavioral responses. Reducing the time interval between detecting the presence of a stimulus and producing a behavioral response is adaptively favorable, for in the natural environment organisms are constantly asked to develop survival strategies under time pressure. Imagine once again our ancestors in one of the many ecological niches they colonized in the late Pleistocene, facing an animal threat, such as a lion. Such a situation requires a very rapid behavioral response priming the animal for fighting or fleeing. Since EM may increase the rapidity of an organism's behavioral response, as we discussed above, it might well be an adaptation.

Furthermore, by integrating parts of the environment in the cognitive loop, the extended mind is supposed to provide the organism with a better interaction with its own ecological niche. In other words, being the environment itself in the cognitive loop, the organism-world coupling is supposed to be more successful. Take for example fish locomotion. Many fishes of the order Perciformes, such as the yellow-fin tuna, as well as some marine mammals such as dolphins, can swim at speeds that can exceed that of gravity by about 20 times. Moreover, they can reverse direction without slowing down and with a turning radius only 10 to 30% of the length of their bodies. Although they basically depend on this locomotive speed and aquatic agility for their survival, these adaptive features do not rest, as one would think, on piscine propulsion nor on purely internal cognitive strategies. Muscular and fin system, indeed, are insufficient both for their flexibility and propelling power to explain such astonishing behavioral properties: it has been calculated, for instance, that dolphins are too weak by a factor of about 7, to attain such speeds. In addition, would the fish exclusively rely on its internal resources to achieve these records, then the expenditure of energy would be too high for its metabolism. This gives rise to a paradox, known as the Gray's paradox, whose inescapable solution is that there are flow mechanisms at work around the body of the moving dolphin that lower its drag by a factor of 7. In other words, to achieve their behavioral records dolphins massively exploit the aquatic environment around them, off-loading to the water the exercise of some functions that are intuitively supposed to be intrinsic of their bodies. Swimming at 50 knots and reversing with a turning radio of 25%, therefore, are records that the dolphin can only obtain through integrated coupling between itself and the aquatic environment. By means of this off-load aquatic animals augment their adaptive coefficient not only in consequence of the improvement of their skills (such as speed and agility), but in consequence of the refinement of coupling with their own environment too. By integrating external features such as water pressure, salinity gradient, fluid temperature, current intensity and direction, aquatic animals also refine their interaction with the ecological niche where they live.

Successful coupling with the environment requires deep (personal and sub-personal) knowledge of that environment. In force of this knowledge, they get more capable to develop successful survival strategies not only in the standard situation, but also when dangerous changes in the environment occur. It has been noted, for instance, that in case a threat in the environment occurs, such as a crude oil spill in the ocean, fishes and dolphins can detect it before the oil invades their optical space.

Although further studies are required, an elegant explanation of this phenomenon is to assume that fishes and dolphins, in force of their well-optimized coupling with and knowledge of the aquatic environment, can detect physical changes in the water, such as in the hydrocarbon level or in the water pressure, and correlate them to a life-threatening danger before to see the oil, similarly as how they can avoid collisions with boats by rapidly detecting a change in current direction and intensity. In a very similar way honey bees, insofar as they integrate external objects such as the hive and the sun position in their cognitive loop, display an astonishing skill at sense of direction and sense of time. It is thus probable for evolution to have primed individuals to succeed in coupling with external objects to better interact with their environment.

4. EMH AND THE SYSTEMATICS OF THE SCIENCES

The evolutionary argument sketched above lays itself open to the criticism that it does not empirically prove EMH to be true but simply corroborates it based on evolutionary theory.

However, Richard Dawkins and other authors have argued that it is conceptually improbable to find a kind of any empirical evidence or experimental design that could uncontroversially prove that the mind-environment couplings responsible for this presumably adaptive off-load of cognitive processes outside the organism's nervous system, should count *de facto* as extensions (i.e. constitutive components) of the mind itself. Determining whether they count as extensions or not, does not seem to be an issue that can be assessed with empirical observation, but rather seems to be a matter of conceptual definitions. It is worth to point out that conceding this does not imply to maintain that the attribute 'extended' referred to the mind is arbitrary after all. Not being subjected to ultimate empirical proof is a condition that scientific propositions may incur at a high level of abstraction. However this does not make the choice among them a matter of mere subjective taste.

Rather, according to the basic principles of theory choice in science, in absence of ultimate empirical proof, the justification of a given hypothesis depends on two properties: 1) the theoretical parsimony of the hypothesis (law of parsimony); 2) the grade of compatibility of the hypothesis with the overall scientific paradigm (law of conservatism). My claim here is that, considering the external off-loading of (phases of) internal processes as extensions of the mind is heuristically suitable for it meets both principles for theory choice in science, given the absence of ultimate empirical proof. I will therefore commit myself to the claim that EMH is more suitable than its antagonist hypothesis.

First of all, EMH is more suitable than unextended accounts of cognitions because it makes fewer assumptions and thereby offers the simplest explanation for mind-environment coupling. And other things being equal, a simpler explanation is better than a more complex one. EMH, in fact, displays the explanatory advantage of admitting only one system for the execution of an environment-involving cognitive function. Whilst traditional views attribute the execution of a given function, say the recall of a phone number stored on the iPhone, to two different decoupled systems, i.e. the mind and the iPhone, EMH attributes this function to the sole integrated system composed of the active coupling of the mind with the iPhone. This assumption is particularly valuable in behavioral perspective, for it allows to avoid decoupled explanations for complex adaptive behavior. For example it allows to attribute the honey bee waggle dance only to the extended mind of the bee instead of to the bee's unextended mind, the hive and the sun position.

Secondly, EMH is more suitable than its antagonist hypotheses because it is more compatible with the current scientific systematics. Indeed, as the evolutionary argument should have shown, EMH is more compatible with the default description of adaptive features. In fact, insofar as it is self-evident for evolutionary biologists to consider adaptations such as lactose tolerance as adaptations of the digestive system, and bipedalism as an adaptation of the muskulo-skeletal system, then there seem to be no impediment for considering the ability to performing environment-involving functional coupling as an adaptation of the mind. But if the adaptation of performing environment-involving functional coupling is a constitutive feature of the mind, then all components concurring in the realization of this function must be constitutive of the mind. Translated into the EMH vocabulary: the mind extends to them. Evolutionary biology, in other words, presupposes the existence of adaptive functions and matching systems executing such functions. Systems executing adaptive functions are considered holistically, regardless either of whether their constituents are all internal or whether some of them reside externally.

Furthermore, EMH is explanatory more compatible with the general scientific paradigm because it provides a description of cognitive faculties of living systems at the functional level. Explanations in ethology and behavioral psychology, for instance, use to attribute cognitive behavioral patterns, such as ant spatial cognition, chimpanzee's rule-learning and octopus' tool use, to the organism performing such behavioral patterns: respectively the ant, the chimpanzee and the octopus. There seems to be no reason then, not to attribute the cognitive behavioral pattern of integrating objects in the environment within the cognitive loop of the organism, to the organism itself. But attributing a cognitive function to a certain organism implies by definition to attribute it to the mind of that organism. And attributing processes involving mind-environment functional couplings to the mind of the organism implies ceteris paribus a commitment to the idea that the mind may extend to the environment. Denying EMH, therefore, would contravene the principle of conservatism in the strong sense that it would bring up the default description of behavioral patterns at the functional level. For example, it would imply either not to consider the honey bee waggle dance as a behavioral pattern of the bee or even not to consider it a behavioral pattern at all. Neither of these options seems to be heuristically attractive.

CONCLUSION

This paper attempted to historically situate the Extended Mind Hypothesis (EMH) in a broader landscape of philosophical and scientific ideas and theories about the mind. Furthermore, it presented an argument in support of EMH. The argument consisted of two sub-arguments, an evolutionary and a systematic one. The evolutionary sub-argument advanced that there is good empirical and analytic reason to suppose that the mind is designed by natural selection to extend. The systematic argument stated that EMH is more consistent with the current systematics of science and scientific taxonomy as it better complies with the principles of parsimony and conservatism. In force of the connection between these sub-arguments, I have concluded that EMH is epistemically advantageous compared to purely internalist accounts of cognition.

BIBLIOGRAPHY

Adams, F., and Aizawa, K. 2001. The bounds of cognition. *Philosophical Psychology*, 14, 43-64.

Adams, F., and Aizawa, K. 2009. "Why the Mind is still in the Head". In *The Cambridge Handbook of Situated Cognition*, edited by A. Murat A. & P. Robbins, 78-95, Cambridge University Press.

Blaauw, G. A., and Brooks, Jr., Frederick P. 1997. *Computer Architecture: Concepts and Evolution*, 489–493. Addison-Wesley.
Bullington, J. (2013). "The lived body". In *The expression of the psychosomatic body from a phenomenological perspectiv*e, 19-37, Springer.
Carpenter, T.P, Fennema, E., Franke, M.L.L., Levi, L., and Empson, S. B. 1999. *Children's Mathematics - Cognitively Guided Instruction*, Heinemann; 1 edition.
Carruthers, P. 2002. The cognitive Functions of Language, *Behavioral and Brain Sciences*, 25(6):657-74.
Clark, A. 1998. "Magic Words: how Language augments human Computation". In *Language and Thought*, edited by P. Carruthers and J. Boucher, Cambridge University Press.
Cowan N., 1998. What is More Explanatory, Processing Capacity or Processing Speed? *Behavioral and Brain Sciences*, 21 (6):835-836
Dartnall, T. 2005. Does the World Leak Into the Mind? Active Externalism, „Internalism", and Epistemology, *Cognitive Science,* 29:135-43.
Dascal T. 2002. Language as Cognitive Technology, *International Journal of Cognition and Technology*, 1 (1):35-61.
Deacon, T. 1997. *The Symbolic Species: The co evolution of language and the human brain*, London: Penguin.
Dehaene, S. 1999. *The Number Sense*, Oxford University Press.
Dunbar, R.I.M. 1996. *Grooming, Gossip and the Evolution of Language*. London: Faber and Faber.
Gallistel, R. and Gelman, R. 1992. Preverbal and verbal counting and computation. *Cognition,* Aug; 44(1-2):43-74.
Geary David, C. 2005. *The Origin of Mind: Evolution of Brain, Cognition, and General Intelligence*, American Psychological Assonciation.
Ginsburg, H.P. 1989. Children's Arithmetic, 2nd edition, Pro-ed, Austin, TX.
Kirsh D. & Maglio P. 1994. On distinguishing Epistemic from Pragmatic Action*,* *Cognitive Science*, 18:513-49.
Li, P., & Gleitman, L. 2002. Turning the Tables: Language and Spatial Reasoning. *Cognition*, 83.
Lupyan, G. 2012. Linguistically Modulated Perception and Cognition: the Label-feedback Hypothesis. *Frontiers in Psychology*, 3:54.
Merleau-Ponty, M. 1945. *Phénoménologie de la perception.* Éditions Gallimard, Paris; English translation by C. Smith, 1962, *Phenomenology of perception*, Routledge and Kegan Paul, London.
Nietzsche, F. W. 1914. *The Complete Works of Friedrich Nietzsche: Thus Spake Zarathustra*, Vol. 11: TN Foulis.
Piccinini, G., & Scarantino, A. 2010. Computation Vs. Information Processing: Why Their Difference Matters to Cognitive Science, *Studies in History and Philosophy of Science Part A*, 41 (3):237-246.
Pinker, S. 1994. *The Language Instinct: How the Mind Creates Language*, Harper.
Pinker, S., & Bloom P. 1990. Natural Language and Natural Selection, *Behavioral and Brain Sciences,* 13.4: 707-726.
Seeley, T.D. 1995. *The Wisdom of the Hive*. Cambridge, MA: Harvard University Press.

Sneddon, A. 2002. Towards Externalist Psychopathology, *Philosophical Psychology*, 15 (3):297-316

Streidter, G.F. 2005. *Principles of Brain Evolution*, Sinauer Associates, Sunderland, MA.

Teeling, E.C., Scally, M., Kao, D.J., Romagnoli, M.L., Springer, M.S., Stanhope, M.J. 2000. Molecular Evidence regarding the Origin of Echolocation and Flight in Bats. *Nature,* 403: 188–192.

von Frisch, K. 1967. *The Dance Language and Orientation of Bees*. Cambridge, MA: Harvard University Press.

Wilson, R. A., and Foglia, L. 2011. Embodied cognition, *Stanford Encyclopedia of Philosophy*.