Extended Knowledge and Causal Dependence

Esteban Céspedes[1]

Abstract: *One of the principal presuppositions in the extended mind account of Clark and Chalmers establishes that extended and non-extended cognitive systems have somehow the same structure and that the distinctions between them can only be superficial. In contrast, this work presents some arguments for the idea that it is possible to find fundamental differences between both, mainly on the basis that a criterion that does not include the notion of knowledge is not strong enough to define cognitive processes (section 2). A brief analysis on the non-transitivity of trust (section 1) and the notion of causal dependence between information and cognitive systems (section 5) might be helpful to support this position. It will be argued (section 4) that the counterfactual block which supports the extended mind building does not seem to be firm.*

1. Four Features

One of the examples the extended mind account is based on could be presented in a summarized version as follows. Inga wants to go to the museum, she recalls where it is and then she goes there. Otto (who has long term memory problems) wants to go to the museum, he reads his notebook where he wrote down last week where the museum was and then he goes to the museum. The extended mind account argues that both cases are fundamentally the same. "Certainly, insofar as beliefs and desires are characterized by their explanatory roles, Otto's and Inga's cases seem to be on a par: the essential causal dynamics of the two cases mirror each other precisely."[2]

In order to say that both cases are the same, there are criteria under which both cases might be analyzed. There are four main features a cognitive system has when a belief takes place, according to the definition of the extended mind

[1] I am indebted to André Fuhrmann for fruitful and guiding discussions that helped me to arrive at some answers in this topic and to the Konrad Adenauer Foundation for supporting this investigation.

[2] A. Clark and D. Chalmers, "The extended mind," *Analysis*, 58 (1998), p. 13.

Iris, ISSN 2036-3257, III 6 October 2011, p. 57-x

account.[3] Firstly, the extension of a cognitive process should be a constant in the history of the main cognitive system. In the example, Otto's notebook is constantly present in his life, there is not a single period of time in which this extension is absent. Secondly, the information managed by the external component is directly available without problems to the main system. That excludes cases where Otto and his notebook exist constantly and permanently as a system, but they do not have any connection that permits information transference. Thirdly, when the information contained in the extension is communicated to the main cognitive system, it is immediately endorsed and accepted by the latter. That permits that Otto accepts what is written on his notebook when he reads it. And fourthly, there is a past-endorsement of the information by the main cognitive system and that endorsement produces the reception of information by the extension of the system in the first place. This last point may be arguable and it is precisely supposing the absence of this criterion, where the problem with this account might lie on.

Suppose a case where the information contained in the external component of the system is not transmitted by the main system, but by another system that is external to both. Let's think of a friend of Otto, who is always helping him and sometimes writes down the information in Otto's notebook that he thinks might be important to him. Such situations need, within the extended mind account, a great amount of trust, reliance and accessibility, although the permanent accessibility is not necessary in this example, for Otto only puts his trust and reliance in his friend's beliefs and what he thinks that might be important to him. In this case, the main system forms again an extended system after including the external components, but the information does not only flow between them. Otto always takes a look at his notebook before answering a question. He would not say that he did not know the answer, if he had not seen it yet. But this time, he does not know with certainty if he or his friend wrote the information he is reading. He will just use that information as if he had written it down and will endorse it immediately after seeing it, which agrees with the third condition for the extended cognitive system.

However, it does not seem that this process completely lacks important and profound differences with cognitive processes including memory. These cases describe complete extended mind systems (in absence of the past-endorsement criterion), but they do not have the essential properties of a cognitive process that might include processes of memory and belief. After all, the extended mind hypothesis does not require that the recall made by Inga is exactly identical with the recalling made by Otto, but that both are processes

[3] Ibid., p. 17.

that are similar enough with respect to their cognitive features. The only differences they might have is as lack of relevance. "The differences between Otto's case and Inga's are striking, but they are superficial."[4] But this case shows an example where the process made by Otto is essentially different from the process made by Inga when they recall something. Otto just believes in his notebook without neither criticizing the information it contains, nor its truth, while Inga has some kind of judgment about the belief she recalls. She has at least some degree of belief or some considerations about the probability of that belief. Otto believes somehow with certainty in the information contained in the notebook, although that certainty is not well founded.

This problem may raise the question about the transitivity of the extended cognitive systems. How far is capable an extended system to be extended? Suppose that a main system assumes information from its extended parts, information that comes from another system, which was produced after having been received by a third different system. Is the last system of the chain part of the whole cognitive system? I cannot see how the extended mind account might answer negatively to this question.

The fundamental supposition of this account is not that every cognitive process possesses the features of an extended process, but that there are cognitive processes that are not in the head and therefore imply external processes that might be included in the complete cognitive system.[5] Nevertheless, there are plenty of good reasons for thinking that the transmission of information does not happen always in a transitive way. Suppose that this case is true:

(1) (If S_1 believes P, then S_2 believes P) & (If S_2 believes P, then S_3 believes P)

Is it true that if S_1 believes that P, then S_3 believes that P? It seems not; not as a rule, at least. This idea is supported by the studies of Christianson and Harbinson,[6] who have shown that the notion of trust is not transitive. Any process of cognition is under certain rules, certain regularities and laws. The case of an extended mind might have a very significant regularity that permits the system to believe every piece of information contained in the extension. This regularity might be, for example, that "For every proposition that is written in the extension system, the main system endorses that proposition." In other words, the only criterion for believing something, in a case where the main system only answers questions after checking the extension system,

[4] Ibid., p. 14.
[5] Ibid., p. 8.
[6] B. Christianson and W. S. Harbinson, "Why isn't Trust Transitive?," in M. Lomas (ed.), *Proceedings of the International Workshop on Security Protocols*, London: Springer, 1997, pp. 171-176.

is to endorse what is written in the extension system. The only truth-maker of those propositions is based on copying the information of the notebook.

The extended mind theorist might argue that there are no fundamental differences between those kinds of regularities and the regularities that govern the mental processes happening inside the brain. He might say, for example, that if *S* believed *P* yesterday, then *S* will believe *P* today (if he does not approve any proposition in that period that contradicts *P*.) But that case is fundamentally different from the case of the extended mind, for those are logical rules, and the extended mind does not care if the information that the extension system is providing in that precise moment is in contradiction with the information it provided the last time. In fact, the extended mind could have a component that is inconsistent, and that does not prevent the main system to endorse that belief.[7] Someone could say that non-extended minds often believe things that contradict each other and that those minds revise a belief in the same way Otto would revise it, if the set of sentences written in his notebook lacked consistency.

But what provides those regularities? I would say that Otto's mind does, the main cognitive system, in other words. Suppose that Otto does not always remember some rules of logic, like the law of non-contradiction, and that he always carries a logic book with him. Would it also be a case of extended mind? He would be somehow able to check the information on the book. I would say that the belief in that system is essentially different from a non-extended belief.

It could be argued again that the neurons are governed by regularities in the same manner that an extended mind is governed by the rule that says that everything written in the extension is also believed and endorsed, and that the neurons are also governed by chemical and physical regularities. Nevertheless, those regularities have the form of natural laws, what makes them something else. These are not mere contingent regularities, like the self-imposed rules Otto has to act with.

2. Extended Knowledge

The extended mind is usually evaluated by the capacities of belief. Is Inga's belief fundamentally different from Otto's belief? If those differences are present, superficiality is everywhere. I think that this assumption might be arguable. Nevertheless, if one considers a criterion of *knowledge* rather than

[7] This problem is based on a coherentist notion of a belief state. The questions related to foundationalism are also going to be discussed here, together with the notion of knowledge.

a belief-based criterion, the results might be more different. Let us define knowledge classically, saying that it is justified true belief. Since Gettier problem this account has experienced many changes, though, but for my purposes it can be considered as plausible.[8] In a different kind of approach, Levi argued that the task of epistemology is to improve knowledge rather than to search for its justification.[9] Gärdenfors has presented in a clear way, following that distinction, the contrast between foundation and coherence theories of epistemic states.[10]

Suppose that Inga *knows* where the museum is and she has therefore support and a justification for that belief, for example, she was there last week. Anyway, the form in which she had justified her belief about the museum is not important now. Otto may also answer, using his notebook, to the question where the museum is located. After reading a sentence about the museum written on his notebook, it is not so doubtful that he will believe the sentence that describes where the museum is. But does he *know* that? Probably Otto would say that he knew that, because he had written it down earlier and he believes in what he has written (and in what his friend has written).

Anyhow, every case of extended mind requires a high amount of trust, some kind of tolerance about the truth of beliefs that are supported on the external system. Not only socially extended cognition requires a trusty mood, after all. Someone might reply that actually the system formed by Otto and his notebook is the one which knows where the museum is located, but that would imply that the role the notebook plays in this cognitive process is as important as the role Otto plays. That is hard to accept. One of the fundamental notions of this account is that cognition is *extended*, that there is a main part to which an external *component* is added.[11] Even if one accepts that the couple Otto-notebook knows that, the knowledge of that couple would be essentially different from the knowledge Inga has about the same proposition. If someone asked Inga to justify her belief, she would be able to make a backtracking (and perhaps somehow complex) explanation of why she knew that. Otto would maybe just say that he believed unconditionally in what appeared in his notebook, perhaps because he wrote it down or

[8] E. L. Gettier, "Is Justified True Belief Knowledge?," *Analysis*, 23 (1963), pp. 121-123.

[9] I. Levi, *The Enterprise of Knowledge: Essay on Knowledge, Credal Probability and Chance,* Cambridge, MA: MIT Press, 1980.

[10] P. Gärdenfors, "The Dynamics of Belief Systems: Foundations vs. Coherence Theories," in C. Bicchieri (ed.), *Knowledge, Belief and Strategic Interaction*, Cambridge: Cambridge University Press, 1992.

[11] On one side, the extended mind hypothesis supposes an identity between the cognitive agent and the extended system taken as a couple. But on the other side, it establishes an *extension* of something that might be perfectly considered as a main system.

because his friend did. But that dogmatic belief does not seem to be knowledge at all.

That could be replied by saying that a non-extended mind is in many cases dogmatic too, for example, after watching the news or reading a scientific journal. But at least the propositions that put that belief in doubt or the change of those beliefs might be justified. Besides, one could believe that the museum is located in a certain place, recognising at the same time that it might be slightly probable that the proposition is not true or that the justification is not valid. But in case of an extended cognitive system (like Otto's), the main system cannot have any degree of belief about the proposition different from the higher; the belief must be accepted in a dogmatic way.

The notion of belief is just not strong enough to explain a cognitive system. The extended mind shall pass the test of knowledge if it wants to be argued that there are no essential differences between extended and non-extended systems. That seems to be very difficult. How strong is Otto's capacity of refuting someone's believes? Very low, I guess. Suppose that a person asks Otto if it is true that there are no museums in his city and he answers negatively by saying that he read it on his notebook. That would not be acceptable as a concluding refutation.

The extended mind theorist might say that Otto also has the information that is necessary to guide that person to one of the museums and that would undoubtedly be an acceptable refutation. I think that the problem here is not whether an extended mind is more or less convincing than a non-extended mind, but whether some mind knows what it believes and whether it can justify (or refute) that belief. And if you have the case, where a mind can justify its beliefs, then the question is whether the *whole* extended mind is justifying that belief or just some part of it. Is just Otto justifying his belief of the location of the museum or is the couple Otto-notebook justifying it? I think that there are better reasons to think that just Otto is justifying that and, furthermore, the cognitive act of justifying something is not independent of the belief that is being justified. Somehow Otto first believes what is written in his notebook and then he starts doing the cognitive processes.

This means that Otto's beliefs are neither inside his notebook, nor is the notebook identical with (or a constitutive part of) his memory. Therefore there is no rational analogy between the memory of a non-extended mind (or a part of it) and the external component of an extended mind, whose function it is to provide information that the main system is going to believe later. Many problems arise after asking those questions about the differences between Otto's functions and the notebook's functions. The next section is related in some points to these problems.

3. Main Systems and Persons

There are close connections between the notion of the mind and the notion of a person, at least when speaking about the human mind. It would not be rational to establish that a mind is identical with the person that is related to that mind, but it would be more than natural to say that having a mind is necessary for being a person. Clearly, there are cases where cognitive processes are not in their optimal performance, but that does not mean that the person is not there. Nevertheless, it is not irrational to think that if the mind completely ceases its activities, the person (or the essential properties of the person) will also cease being. In other words, the relation is not symmetric; the person could completely cease his or her activities and the mind could continue operating.

In a recent article, Baker argues that the notion of enduring persons is crucial to understand the problems of the extended mind account. Furthermore, if there is some extension to a certain mind, this extension would not be a part of it, the mind would not become the couple mind-extension.

> We are still agents and subjects of experience, not mere systems or components of systems. Cognitive processing does loop out into the world, but processing does not stand on its own. It requires an entity that is doing the processing. Processing does not perceive or act on the world; we do. Brains do the processing that enables us to perceive and act on the world, but the entities who act on the world are not brains – they are agents.[12]

This view is not going to be accepted here as a whole, but this argument seems to be very strong in order to understand why there is a main cognitive system that, even when the total system is an extended one, is better described as a mind than as the sum of the interactions that it might have with other systems. When Otto has to justify his belief about the location of the museum, is Otto as an enduring agent with his cognitive capacities the only one who justifies his belief and justifies why he has to believe what is written in his notebook, or is it the couple Otto-notebook that justifies it? Even if the notebook contained all the information that is necessary to justify his belief about something or to justify the rule that prescribes that he has to believe what is written there, it would not be reasonable to think that the complete extended system is doing the justification.

Suppose that, in that case, the couple Otto-notebook is justifying some belief. Even in that scenario, Otto's friend might be included in the example

[12] L. R. Baker, "Persons and the extended-mind thesis," *Zygon*, 44 (2009), p. 646.

and (ignoring the fourth feature) it might be asked, whether the friend plays an important role. If the set formed by Otto and his notebook is taken as an extended cognitive system, is there any reason to exclude his friend as part of that system? And if he is part of the system, he also constitutes the system. Is the cognitive process of knowing where the museum is located also performed by Otto's friend? If that is true, is it the friend as agent who constitutes the extended system of Otto or is it just his memory?

In the extended mind account, all those things must be considered as part of the extended mind, and that complete system would not have fundamental differences with a non-extended system involved in the same kind of cognitive process. That does not sound rational, although it is a conclusion of the extended mind hypothesis. But the problem does not seem to lie on the mereological distribution of a system,[13] but on the notion of spatial layout that is related to cognitive processes, which will be considered in the next section.

4. Mind and Skull

One of the central theses of the extended mind account is that not every cognitive process is performed inside the head.[14] In other words, it tries to refute the hypothesis that all the cognitive processes are performed in a certain spatial region. That hypothesis could be analysed like this, supposing that *c* is a mental process, *S* a subject, and *h* is his head (or some spatial region inside his skull).

(2) $\forall S \forall c \exists h$ (*c* is a mental process of *S*) $\rightarrow$ (*c* is performed inside the spatial region of *h*)

And the thesis of the extended mind would be this:

(3) $\exists S \exists c \exists h$ (c is a mental process of S) & $\sim$(*c* is performed inside the spatial region of *h*)

In the case of extended minds, the external features (or extensions) are important in some *active* way, they are vehicles of the cognitive process. It

[13] See, for example, the treatment of the mereological fallacy in neuroscience, analyzed by M. Bennett and P. Hacker, *Philosophical Foundations of Neuroscience*, Oxford: Blackwell, 2003, p. 73. This is a misconception that arises when the functions of the whole are attributed to a part, like some particular area of the brain.

[14] Clark and Chalmers, "The Extended Mind," p. 8.

could be said that they play certain causal roles, without which the cognitive process might not have been produced, a topic which is going to be discussed in the next section. At least two important elements of the argument have the form of a counterfactual conditional.[15] Firstly, if the external process had been performed inside the head, it would have been rational to consider it as a cognitive process. Secondly, if the external process had not existed, the cognitive process would not have existed. The first counterfactual reasoning might have this form, considering that *e* is an external process which can consitute an extended cognitive process, like Otto's notebook:

(4) (*e* occurs in *h*) €→ (*e* is a cognitive process)

This is some kind of presupposition connected to the realm of common sense and is fundamental in the argument of the extended mind account. Looking at the example, if the information of the notebook had been available to Otto inside his skull, it would have been rational to consider that belief as a complete cognitive process. Suppose that somehow (the futurist geek details are not needed here) the belief about the location of the museum is put into Otto's skull, omitting the sentences about the justification of that belief. The same sentence that was written on the notebook is transported to Otto's collection of propositions in his memory. If someone asks Otto whether he believes that, he will answer that he does. But what would he answer, if someone asked whether he *knew* that? Brainwashing is not knowledge and the effects of brainwashing are not justifiable beliefs.

The extended mind theorist might reply that his account is not about content, but about the activity and the vehicle of the cognitive processes. Furthermore, the extended mind account does not seem to contradict internalism in terms of content, like Bartlett explained recently.[16] But even in that case, a belief's justification must not only be about content; its vehicles are also important as justifications. I think that going into these kinds of issues would only lead to strong reductionist views about the mind and to the problem, whether the mind is able to be extended not only outside but also inside the skull. There are other problematic issues about the causal dependence in the structure of both processes taken as vehicles. The discussion about the second counterfactual reasoning mentioned above might be more interesting.

[15] Ibid., pp. 8-9.

[16] G. Bartlett, "Wither Internalism? How Internalists should Respond to the Extended Mind Hypothesis," *Metaphilosophy*, 39 (2008), pp. 163-184.

5. Causal Dependence

It was suggested above that other assumptions of the extended mind account were about the impossibility of elimination of the external process. If the external process *e* had not occurred, the cognitive process *c* would not have existed.[17]

(5) ~e ∈→ ~c

It is easy to realize that this conditional has the same form as the notion of causal dependence in the sense that counterfactual theories have developed it since Lewis.[18] In some way, the notebook plays a fundamental causal role in Otto's belief. If he did not have his notebook, he would not have the belief about the location of the museum. This partly seems to explain why the first feature of constancy is needed in order to define a mental state of belief and to sustain that this account is called *active externalism*, because the extensions play a causal role in the cognitive process. However, the fundamental characteristic is that the external feature cannot be eliminated, if the cognitive process is wanted to be performed:

> The external features in a coupled system play an ineliminable role – if we retain internal structure but change the external features, behavior may change completely. The external features here are just as causally relevant as typical internal features of the brain.[19]

Clearly, if someone (let us say Otto's friend) changes some sentences in Otto's notebook in order to make him believe something wrong, he will change his beliefs accordingly and his behavior may change too. But what would happen if his friend explained him the joke, immediately after having changed the sentence? He would probably think that he better erased the note and changed it. The notebook (*n*) does not include that process; some main feature of Otto's mind is playing a more crucial role here. Let □ be a binary relation between events, where "*c* □ *e*" stands for "*c* causes *e*".

(6) (Friend modifies *n*) □ (*n* has the information that *P*) □ (Otto believes that *P*)

(Friend modifies *n*) □ (Otto believes that ~*P*)

[17] Clark and Chalmers, "The Extended Mind," p. 9.
[18] D. Lewis, "Causation," *The Journal of Philosophy*, 70 (1973), pp. 556-567.
[19] Clark and Chalmers, "The Extended Mind," p. 9.

(Otto modifies *n*) □ (*n* has the information that ~*P*)

The fact that something plays an important causal role in a cognitive process does not mean that it should be included as a part of that process. Oxygen molecules also play an ineliminable role in every belief, but it would not be rational to think that they are part of an extended part of it. It is clear, though, that the constancy of the external feature is just a necessary (not a sufficient) condition for belief.

Since causal roles are mentioned, maybe the notion of *effect* should be considered more closely. No one would argue against the claim that believing that *p* is the case is caused by something. In the situation of Inga, she believes that *p* is the case for reasons that differ from the reasons Otto has when he believes that *p* is the case. The direct cause of Inga's belief might be some material feature inside the skull, while the direct cause of Otto's belief is the external feature, also material, the writings inside the notebook. What is the cause of those writings? Clearly, Otto; there is some kind of causal loop in that process, an element that distinguishes this account from other kinds of content externalisms.[20] But it must be detected that there is also a loop inside Inga's skull and that is precisely the point the extended mind theorist wants to establish! Let us say that *b* stands for an occurring brain process.

(7) (Otto believes$_{t1}$ that P & writes$_{t1}$ '*P*') & (Otto reads$_{t2}$ '*P*' & believes$_{t2}$ that *P*)
(8) (Inga believes$_{t1}$ that *P* & b_{t1}) & (b_{t2} & Inga believes$_{t2}$ that *P*)

It does not matter now what the causes of their beliefs at instant t_1 are, that might just require a normal externalist analysis. The important processes are the writing, the reading and the brain processes in between. Those events are the causes of the loop and the reason why the cognitive processes in both cases do not have fundamental differences for the extended mind account. But unfortunately, there are essential differences between the whole cognitive process in situation (7) and the one in situation (8). Whereas the brain processes inside Inga's skull do not imply mental processes (there are just firing neurons and synapses), the writing and the reading necessarily imply mental processes in between.

I do not want to argue here that extended mind account is wrong because it establishes that not every cognitive process is performed inside a certain spatial region. In order to refute that, someone would have to be able to show

[20] See H. Putnam, "The meaning of 'meaning,'" in K. Gunderson (ed.), *Language, Mind, and Knowledge*, Minneapolis: University of Minnesota Press, 1975; and T. Burge, "Individualism and the mental," *Midwest Studies In Philosophy*, 4 (1979), pp. 73-121.

that *every* cognitive process does actually happen inside that region, but trying to find the place of mental processes is an irrelevant task in both cases. I just argue that the cases that seem to be indistinguishable to the extended mind theorist are in fact fundamentally different. Furthermore, *that* is the reason why active externalism does not explain correctly what the mind is, and not because there is some way of showing that every cognitive process is internal. Actually, there is no rational way to show that. Thus, I do not agree that avoiding the issue of where the mental states and processes are realized is a way of avoiding the challenge of extended cognition.

Supposing that every belief is an effect and that it depends causally on other events, there is a criterion that was mentioned above, in order to which extended and non-extended mental processes might be distinguished. How does someone find the cause of an event, for example, of an explosion? He establishes the effect, together with other hypotheses and conditions, and deduces then the cause.[21] The most important process in recognizing and explaining causal dependence is not the deduction of the effect from the cause, together with laws and initial conditions, but the possible *refutation* of the effect (together with laws and conditions) by a new possible cause.

Let us consider Otto's beliefs about the location of the museum as an effect. In order to make a refutation of that part of his belief state, better replacing reasons and justifications of that belief and of belief changes must be found. An important role that cognitive processes play is that of refutation or, in other cases, of justification of beliefs, of knowledge. There is a certain dogmatism occurring in Otto's supposed extended cognitive process, which is one of the most fundamental differences between his belief about the museum and Inga's. She could easily make a reasonable and complex backtracking justification of her belief, while Otto could just say that he must believe in what is written in his notebook.

Again, it could be argued against this distinction, that Inga could perfectly fail in making a justification of a belief she possesses, if she does not remember how she came up to believe that. Even in that case, if the extended mind theorist says that the action of going to the museum is in both cases an effect of wanting to go and believing where the museum is,[22] the precedent cognitive process is not the same. It could be accepted that both cases are somehow analogous (the notebook plays the role of the memory), but that relation would not be stronger than a metaphor, never being two cases of the same kind.

[21] This analysis has some fundamental differences with the covering law model of explanation, in which the sentence describing the cause is part of the premises. Here, the cause is deduced partly from the effect.

[22] Clark and Chalmers, "The extended mind," p. 13.

6. Concluding Remarks

The arguments that might support essential (non-superficial) differences between cases of extended cognition and non-extended cognition are, at least, five. One of them (section 2) establishes that an extended mind cannot experience knowledge defined as justified true belief and that, if the aim of the extended mind account is to provide a definition of the concept of a mental process, it must be based on stronger notions than the notion of mere belief. A second argument (section 5) is a result from ignoring the condition that demands past-endorsing by the main cognitive system. This change might lead to absurdity within extended mind examples, because of the non-transitivity of trust. Another argument (section 1) establishes that a non-extended process is fundamentally governed by natural laws, while an extended cognitive process is governed by contingent regularities. A fourth argument (section 4) argues the falsity of a counterfactual conditional that lies on the basis of the extended mind building, which says that if an external cognitive process (of an extended mind) had occurred inside the head, it would have been considered a cognitive process. And a last argument (section 5) establishes that while Inga's belief is constituted by one simple mental process, Otto's cognitive process is always formed by a different structure, a causal chain of more than one simple mental process.

Esteban Céspedes
Goethe Universität Frankfurt am Main
e.cespedes@stud.uni-frankfurt.de