

The Extension of the Extended Mind

James Brown

Lancaster University

1 Introduction

I have a mind and a body, and it seems sensible to think that the former is located somewhere within the latter (in my head, perhaps). According to the extended mind thesis, however, this is not necessarily the case: under certain circumstances, the mind can wander beyond the boundaries of the body, beyond skin and skull.

This idea first appears in *The Extended Mind* (Clark and Chalmers 1998) where they argue that cognitions and beliefs can reside in processes outside the body, thus the mind 'extends'. If this thesis is accepted, how far should we go? If other mental phenomena can be shown to extend, the thesis itself seems subject to an extension of sorts. Hence the object of this essay: How might the extended mind thesis extend?

I will argue (in section 3) that there are two methods to extend the extended mind thesis: alter the original criteria to accommodate different mental phenomena or find different mental phenomena that fit the original criteria. The second method, however, is suspect. Clark and Chalmers's criteria apply to the functional roles played by cognition and belief so it's hard to imagine other phenomena playing this role which are not themselves instances of cognition and belief. The second method must be handled with care and in section 4 I reject as misguided Colombetti and Roberts' attempt to extend it to emotions. I surmise (in section 5) that attempts to extend the extended mind thesis must invoke the first method.

The Extension of the Extended Mind

Before all of this, let me sketch the extended mind thesis itself

2 The extended mind

The *Parity Principle*:

If as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process. (Clark and Chalmers 1998, 8)

Suppose that two people engage in solving a problem that involves the rotation of shapes displayed on a screen. The first person rotates the shapes mentally, while the second rotates them physically by pressing a button. Clark and Chalmers (1998) (henceforth C&C) argue that for the second person, the button carries out the same function as the mental process by which the first person rotates the shapes. By the *Parity Principle*, therefore, the button is part of the second person's cognitive process. In such cases, the human organism links with an external object, creating a coupled system, or loop, that is itself cognitive. The external object, for the duration that it is in the loop, governs behaviour in the same sort of way as cognition.

C&C consider two possible objections. First, the loop cannot be cognitive because the cognitive is conscious and the conscious does not extend beyond the head. C&C reply that there are states that are both cognitive and unconscious (e.g. the retrieval of memories). Second, the loop cannot be cognitive because the cognitive is bound to the mind while the coupled system is too easily decoupled. C&C respond that the biological brain can occasionally decouple from its capacities (e.g. consider being intoxicated). All coupling requires is that an external component is reliable, that is, generally there when required. Cognition, therefore, can extend.

The Extension of the Extended Mind

C&C take things a step further: they argue that beliefs can extend too. Consider Otto and Inga both of whom hear about an exhibition at the Museum of Modern Art and decide to go see it. Before setting off, Inga recalls that MoMA is on 53rd Street. Inga's belief that MoMA is on 53rd Street was not previously occurrent but neither are most beliefs; rather, it was embedded in memory, waiting to be accessed. Otto suffers from Alzheimer's disease. To help structure his life, he relies on a notebook. On learning some new information, he writes it down and when he needs it he looks it up. Before setting off for MoMA, he looks up in the notebook and sees that MoMA is on 53rd Street. Otto walks to 53rd Street because he wants to go to MoMA and believes that it is on 53rd Street. He has this belief even before consulting the notebook. The notebook plays the same role for Otto as biological memory plays for Inga.

C&C consider an alternative explanation: Otto does not believe that MoMA is on 53rd Street, rather he believes that MoMA is at the address in the notebook. On filing away the notebook, the belief about the location of MoMA disappears. C&C argue this is wrong just as it would be wrong to claim that Inga's belief disappears when she is not conscious of it. In both cases, the information is reliably there when needed. Further, it is available both to consciousness and to guide action in just the way a belief is expected to be. The causal dynamics in each case are analogous. Since beliefs are characterised by explanatory role, the alternative explanation is spuriously complex.

To reject this position one needs to provide a substantial reason to insist that the 'belief role' cannot be played from outside the body, that Otto's and Inga's cases differ in some relevant respect. C&C entertain five such reasons. First, Inga's

The Extension of the Extended Mind

access to information is more reliable. Inga's access, however, is far from perfect: it can be unreliable (e.g. after having too much to drink). Second, Otto's access comes and goes. Still, Inga occasionally 'disconnects' from her biological memory (e.g. after another drinking session). Third, Inga's access is better. Connection quality alone, however, does not make the difference between believing and not believing; what matters is ease of access to information. Fourth, Inga's access is direct (by introspection, perhaps), whereas Otto's is not (his is by perception). Observe that this begs the question. The claim is that Otto's internal processes together with his notebook constitute a single cognitive system. The flow of information between notebook and brain is, in relevant respects, analogous to the flow of information within Inga's brain. Fifth, Otto's belief is not occurrent, but true beliefs are occurrent; therefore Otto does not have the belief in question. Assuming this is true, Inga's belief is ruled out along with Otto's, which is absurd. Thus C&C argue that differences between Otto's case and Inga's are shallow. For Otto, the notebook entries play just the sort of role that beliefs play in guiding behaviour; it supports all the relevant dispositions.

In closing, C&C list the features that make the notion of 'extended belief' applicable to Otto's case: the notebook is constant, the information in it is directly available without difficulty, and retrieved information is automatically endorsed.¹ What is central is a high degree of trust, reliability and accessibility.

3 Extending the extended mind

In this section, we will see that the central commitments of the extended

¹ C&C mention a fourth feature: prior conscious endorsement. I omit discussion of it since C&C doubt its status as a criterion for belief.

The Extension of the Extended Mind

mind thesis yield two methods for extending the thesis to mental phenomena other than cognition or belief. First, it will be helpful to reformulate the thesis in terms of *vehicles*.

Inga's biological memory systems, working together, govern her behaviour in a distinctly belief-like way; it makes sense to ascribe to Inga the belief that MoMA is on 53rd Street. Otto's behaviour, given his condition, is not governed by biological memory systems; rather, it is governed by a biotechnological matrix, one that consists in himself and his notebook. In both cases, the governing is of the same sort; since it is distinctly belief-like in Inga's case, so too is it belief-like in Otto's. It makes just as much sense to ascribe to Otto the belief that MoMA is on 53rd Street as it does to ascribe it to Inga.

There is a striking, albeit ultimately superficial, difference between Inga's case and Otto's. The stuff that constitutes her belief is wrapped up neatly inside her head, while the stuff that constitutes Otto's extends into his environment. In other words, the *vehicles* that constitute Inga's belief are strictly internal whereas those that constitute Otto's are not. The extended mind thesis, it transpires, is actually about extended *vehicles*: a mind, with respect to some mental property (a belief, say), extends *iff* the set of *vehicles* constituting that property extends, that is, are external. It is still unclear, though, precisely what work vehicles do in constituting a mental property, and precisely how they do it.

Clark (2008) explains that systems like Inga's biological memory or the Otto-notebook matrix can have mental states (or mental states can be ascribed to them), mental states have contents, and contents have vehicles. Further, he defines a *vehicle* as an enduring material object or process that plays a 'special' role in

The Extension of the Extended Mind

enabling the system to have a mental state or in making it reasonable to ascribe a mental state to the system. This warrants closer inspection.

A vehicle is a thing had by a content, which is a thing had by a mental state, which in virtue of the vehicle playing a special role, is a thing had by a system.² First then, what is it for a mental state to have a content? Recall that for the information in his notebook to function in a belief-like manner, or play the ‘belief role’, it must be available to Otto’s consciousness. The information, that is, must be available to be had by consciousness. Specifically, the information is had by Otto’s consciousness in a belief-like way. A mental state, therefore, has the content onto which it is directed: Otto’s belief has ‘MoMA is on 53rd Street’.

In both Inga’s and Otto’s case, the content is the proposition ‘MoMA is on 53rd Street’. The vehicle is the material thing or system that delivers that content. For Inga, the vehicle is her biological memory; for Otto, his notebook. The delivering is precisely the role that a vehicle plays in enabling a system to have a mental state. If content goes undelivered, it is unavailable to consciousness. Such content cannot be had by, or be the object of, a belief or indeed any other content-full mental state. So, because ‘MoMA is on 53rd Street’ is delivered by it, the vehicle enables the belief that MoMA is on 53rd Street.

In Inga’s and Otto’s case, the content is the same but the vehicles differ. If anything, it seems that the relation of ‘having’ should go the other way around. It seems that Inga’s biological memory and Otto’s notebook have, tucked away within them, ‘MoMA is on 53rd Street’, but this misconstrues the relation between content and vehicle. It just so happens that in Otto’s case, the vehicle is a notebook,

² Admittedly, this ‘had by’ relation is obscure, but the obscurity is inherited from Clark’s (2008) explanation that systems ‘have’ mental states, which ‘have’ contents, which ‘have’ vehicles.

The Extension of the Extended Mind

but this is unimportant. A vehicle is not to its contents what a notebook is to its entries. The relation is more like the relation between a commuter and the train: a commuter has the train as a means by which to reach her destination; analogously, a content has its vehicle as a way to 'fill' consciousness in a particular way (e.g. belief-like).

Still, how precisely does a vehicle constitute a mental state? A vehicle constitutes a mental state by playing a certain role, a role as a delivery of sorts. The delivery role must satisfy the following conditions which Clark (Clark 2008; 78) formulates as the criteria to be met by nonbiological candidates for inclusion into a cognitive system:

- (1) The resource is reliably available and typically invoked.
- (2) Retrieved information is more or less automatically endorsed; it is deemed about as trustworthy as something retrieved from biological memory.
- (3) Information contained in the resource is easily accessible when required.
(Clark (2008); 78)

In virtue of these conditions a vehicle constitutes a belief. If the vehicle also happens to be external, it constitutes an extended belief.

Now, (1) is clearly a feature that the vehicle itself must have; (2), a feature that the contents, those that have the vehicle in question, must have; and (3), well, it is unclear. It is unclear whether the contents are easily accessible in virtue of some feature they have, say the notebook's entries are very legible, or in virtue of their vehicle. Intuition seems to suggest the latter, that is, the entries in the notebook seem to be easily accessible due to some feature of the notebook (its

The Extension of the Extended Mind

being well organised, perhaps), rather than to some intrinsic feature.³ Thus for current purposes, (3) will be considered a feature that the vehicle must have. For Otto, the notebook indeed has (1) and (3); and the information it delivers, ‘MoMA is on 53rd Street’, has (2). The role played partly by the notebook itself and partly by the information it delivers constitutes an extended belief.

By satisfying the three central criteria, Otto’s notebook is part of his cognitive system. Does this mean, however, that it also part of his belief system? After all, C&C observe that cognitive extension might not entail the extension of a mental state like belief. Tetris players use the button to complete their cognitive task but the button does not satisfy conditions (1), (2) and (3), the conditions for being a belief. Recall that cognitive coupling only requires that an external component is reliable or generally there when required. This echoes feature (1) and perhaps faintly, feature (2). These requirements for cognitive coupling are clearly less stringent than the conjunct of (1), (2) and (3), the conditions for belief. Because beliefs satisfy (1), (2) and (3) they satisfy the requirements for cognitive coupling but the converse is not necessarily true so, for example, a reliable, generally there-when-needed component need not necessarily deliver contents that are automatically endorsed. Thus the criteria given by Clark (2008) are in fact more than enough for cognitive inclusion: they are for inclusion into a belief system.

To summarise: a belief is a mental state whose content is (2) automatically

³ Note that this does not undermine the earlier claim that a vehicle is not to the content it delivers what a notebook is to the entries it contains. The interaction between Otto, his notebook and its entries is an instance of a vehicle delivering content to a system, but this does not mean it cannot hint at a more general relationship.

The Extension of the Extended Mind

endorsed on retrieval; and whose content is delivered by a vehicle that is (1) reliably available, typically invoked and (3) allows contents to be easily accessed as and when required. Of the vehicles that constitute a mental state, if just one is external, the mental state is extended (or if just one can be external, the mental state is extendable). This captures the essence of the extended mind thesis; but what is it to extend this thesis?

C&C themselves, in fact, had already extended their thesis, calling it ‘taking things a step further’ (12): the conditions for cognitive extension, or coupling, are extended to accommodate belief. To extend the extended mind thesis is to seek to accommodate more types of mental states. C&C introduce criteria to accommodate belief before seeing whether an external component can satisfy them. Otto’s notebook, C&C argue, satisfies the criteria; therefore it constitutes an extended mental state.

There is, however, another method by which the thesis might extend. Instead of introducing different criteria, introduce a different mental state, one that satisfies the existing criteria. On the face of it, this method seems to have poor prospects: if the different mental state satisfies the existing criteria but is neither a cognition nor a belief, what is it? This deserves a closer look. In the following section, I will analyze an attempt to extend the extended mind thesis by this second method.

4 A critique: the case for affectivity

Colombetti and Roberts (2014) (henceforth C&R) propose that the arguments for extended cognitive states apply to affectivity: if extended cognition, then extended affectivity. I will try to show that their arguments almost entirely fail. The objection is not that their notion of affectivity is ill-conceived but rather

The Extension of the Extended Mind

that that the original criteria do not apply to their notion of affectivity.

C&R claim that transient cognitive states (like cognitions) and more permanent mental ones (like dispositional beliefs) are of the same type but differ in degree. In particular, the vehicles that constitute both cognitive and mental states are durable and reliable but the vehicles that constitute mental states are more durable and more reliable than those that constitute cognitive states. Arguments for extended cognitive states differ from arguments for the extended mental states. This, C&R explain, is important to keep in mind, as affective phenomena comprise both transient, occurrent states ('episodes') and more permanent ones (like dispositional resentment).

So far, this should all be fairly familiar. As discussed in the previous section, the requirements for inclusion into a belief system are more demanding than those for inclusion into a cognitive system, but they are similar. However, it is important to note that C&R ask more of extended cognitive loops than C&C: not only must affective loops be reliable, or generally there when required, but they must also be self-stimulating and designed, or selected and maintained, for a certain purpose over time.

C&R interpret the *Parity Principle*⁴ as follows:

If some hybrid organism-world system plays roughly the same role as something that we already call cognitive/mental, then we are justified in seeing this system as a cognitive/mental one. (C&R 2014, 1246)

Again, this is nothing new. Observe how the hybrid Otto-notebook system is cognitive/mental precisely because it plays roughly the same role as Inga's

⁴ Compare: "If as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process." (Clark and Chalmers 1998, 8)

The Extension of the Extended Mind

cognitive/mental system.

Before invoking the extended mind arguments, C&R categorize the various phenomena that fall within the affective domain: there are emotions (fear, anger, sadness etc.), which can be occurrent or dispositional; occurrent emotions or emotional ‘episodes’, which involve cognitive evaluations, bodily changes and feelings; moods, which are not directed at specific objects; and sentiments, which are multi-track affective dispositions.

First, we will consider C&R’s argument for extension to occurrent emotions, or emotional episodes. C&R say an emotional episode has cognitive, bodily and phenomenal components. If an emotional episode has but one extended component, then it is extended. C&R then argue that the cognitive component can extend (1257–1258), so too the affective phenomena. However, if it transpires that affective phenomena do not have cognitive components, the original arguments lose their grip and do not apply.

Enter C&R’s first mistake: they proceed to apply the original criteria to the other two components, neither of which are cognitive. First, they claim that the bodily component can function as some kind of extension but if this is true, then not by the original criteria which were not designed with this sort of phenomena in view. Rather, they were designed in view of the functional role of cognition. Recall that C&C spend a lot of energy identifying the functional roles of cognition and belief. The energy is justifiably spent as it yields a set of functional criteria that an external component must satisfy in order to partly constitute a cognitive or belief system. C&R spend no such energy on applying them to the bodily component of emotional episodes but it is of paramount importance to their cause to determine

The Extension of the Extended Mind

what an external component needs to do to be on a functional par with internal bodily components. For this picture to convince us that the extended system is on bodily par with an internal system (the emotions, for instance), the intricacies of the functional role must be specified. C&R make no such specifications. Thus the *Parity Principle* is not well grounded.

C&R's argument that the phenomenal component can extend is founded on a similar omission. For instance, they introduce the example of a grieving jazz saxophonist who, taken together with her saxophone, comprises a self-stimulating loop: the saxophonist's emotional experience governs what she plays which in turn governs her experience. C&R argue that the phenomenal character of the saxophonist's experience extends across her and the saxophone, that it makes sense to ascribe grief or sadness not just to the saxophonist herself but rather to the loop consisting of both her and the saxophone. It is not that the saxophonist feels sad but rather that both she and the saxophone feel sad – the saxophone is a proper part of her sadness. C&R argue that because the loop self-stimulates, the phenomenal character of the saxophonist's experience pervades it but why should this be the case? Granted, self-stimulation is a central feature of cognitive loops but what motivates the assumption that it is also a central feature of phenomenal loops? Is phenomenal character the sort of thing that self-stimulates? If so, phenomenal character, in at least one respect, is functionally similar to cognition, so the (original) arguments for cognitive extension might be applicable. C&R, however, give no reason to hold that phenomenal character is similar to cognition in this respect. Thus the extended cognition arguments provide no basis for arguing that the phenomenal component of emotional episodes can extend. Once again, the

The Extension of the Extended Mind

Parity Principle is of no use unless the function of the internal process (in this case, phenomenal character) is specified.

Of C&R's three arguments for the extension of occurrent affective states, two miss the mark: the original criteria for cognitive extension are dragged, without justification, into the unfamiliar territories of bodily and phenomenal components. What then of their arguments for enduring affective states? C&R introduce the example of Eve who resents things that her parents do. She writes those things down in a diary. On reading the diary, Eve feels her resentment, that is, the diary fuels Eve's resentment. Without it, she is not disposed to manifest parent-directed resentment.

C&R argue that Eve's interaction with her diary satisfies the 'glue and trust' criteria, that is, criteria (1), (2) and (3) (see section 3): the diary, a content-delivering vehicle, is reliably available and allows easy access to the content it delivers, a content which is automatically endorsed. Recall, however, that the 'glue and trust' criteria are designed to determine whether an external component can partly constitute a belief (or cognition). It seems odd, then, that C&R use them to determine whether the Eve's diary can partly constitute her resentment. To pick on just criterion (2): how often is the content of resentment automatically endorsed? Eve resents things that her parents do but does she automatically endorse those things? Perhaps she does, but this does not imply that the content of resentment is always automatically endorsed, or even just endorsed. For instance, it seems reasonable to suppose that something might be resented because it is not endorsed, in which case the content of resentment is most certainly not endorsed.⁵ Even if it

⁵ It seems just as reasonable to suppose that something might be resented because it fails to satisfy the first or third glue and trust criteria. That is, something might be resented because it is

The Extension of the Extended Mind

turns out that the content of resentment has to be endorsed, does this requirement hold for other emotional dispositions? Take fear, for example. Fear is often of the unknown, but an unknown thing, precisely because it is unknown, cannot be endorsed. While the second criteria seems inapplicable to resentment, it is most certainly inapplicable to emotional dispositions in general. Thus the glue and trust criteria seem inapplicable. Unwittingly, C&R show that Eve's diary can be included into her belief system, but they fail to show that this entails inclusion into her emotional system.

C&R do respond to a similar objection, viz. that the supposed examples of extended affectivity really are instances of dispositional beliefs. Their response is that not all of their examples can reduce to beliefs because some of them involve components that have no propositional content. Moods, for example, are not directed at specific objects; they have unclear contents. Therefore, C&R argue, an extended mood cannot reduce to belief.

This response, however, is inadequate. First, it only addresses the objection as it applies to moods. Granted, moods might not reduce to beliefs but what of 'content-full' affective phenomena (e.g. Eve's resentment)? Moreover, the response undermines C&R's application of the glue and trust criteria to 'content-less' affective phenomena. The glue and trust criteria apply to Eve's resentment because it has content but moods are content-less. A content-less phenomenon presumably has no vehicles whose role is to deliver content. How then can the glue and trust criteria apply to them in any respect whatsoever? Two of the criteria must be met by the vehicle of a state and the other by its content. Surely, a non-existent content
unavailable or inaccessible.

The Extension of the Extended Mind

is not automatically endorsed and a non-existent vehicle cannot be reliable or allow easy access to its (non-existent) content. It seems absurd to assume otherwise. To vehicle- and content-less phenomena, the glue and trust criteria are inapplicable.

It might be protested that C&R only claimed that moods have unclear content, not that they have no content. In this case, even if moods are susceptible to the original objection, they just reduce to being instances of belief. If moods have content (clear or unclear) they reduce to being beliefs; if they lack content, the glue and trust criteria do not apply. A dilemma for C&R.

C&R also argue that temperaments, sentiments and character traits can extend, but these arguments are grounded in similar confusion. In their conclusion, C&R maintain that only an overly strict conceptual separation between cognition and emotion supports the claim that one but not the other extends. But we have seen the separation does not need to be overly strict; rather, any separation seems to render the extended mind thesis arguments inapplicable to affectivity.

In closing this section, I want to grant C&R some success in extending affectivity but to a much lesser extent than they presume. Affective phenomena that have a cognitive component can extend but this is merely a rather unexciting corollary of C&C's original arguments. Indeed, arguments for the extension of affectivity need to be built on the notion of parity but before deciding whether an extended process can be on a functional par with an internal process, the functional intricacies of the internal process itself need to be specified. C&R make no such specifications; they do little more than dress belief up in affectivity's clothing.

5 Conclusion

The object of this essay has been to determine how the extended mind thesis

The Extension of the Extended Mind

might be taken further, or how it might extend to accommodate mental phenomena other than cognition or belief. Seemingly, there are two methods: tweak the original criteria for a mental state that they accommodate a different kind of mental phenomenon; or find a different kind of mental phenomena that satisfies the original criteria.

Colombetti and Roberts (2014) invoke this second method, but to little effect. This alone does not invalidate the second method. Recall, though, that the original arguments were constructed to accommodate the functional roles of cognition and belief. It seems likely, then, that any phenomenon that yields entirely to these arguments will be functionally equivalent to either cognition or belief. And here is the problem: any phenomenon that is functionally equivalent to either cognition or belief is either a cognition or a belief. By the second method, therefore, the extended mind thesis extends nowhere.

Fortunately, we have the first method to fall back on. To extend the extended mind to accommodate an affective state, say, first determine the functional role (or functional features) of that state. Then, use the *Parity Principle* to decide whether this role can be played from outside the body. If so, the state in question can extend. Thus the extended mind thesis extends to accommodate that state.

A final thought on the first method of extension. If the roles of cognition and belief are replaced by the role of a different mental phenomenon, does enough of the original thesis remain to justify the claim that it has extended? The *Parity Principle* will survive, but what of Inga and Otto? Without them, the original arguments seem to lose something essential. If the extended mind thesis consists in more than just the *Parity Principle*, then perhaps its purview expires at belief.

The Extension of the Extended Mind

References

- Clark, A. (2008). Supersizing the mind: Embodiment, Action, and Cognitive Extension (76–81). New York: Oxford University Press.
- Clark, A. and Chalmers, D. (1998). ‘The Extended Mind’, Analysis 58, 7–19.
- Colombetti, G. and Roberts, T. (2014). ‘Extending the Extended Mind: the Case for Extended Affectivity’, Philosophical Studies 172, 1243–1263.