

There is no Argument that the Mind Extends

On the basis of two argumentative examples plus their 'parity principle', Clark and Chalmers argue that mental states like beliefs can extend into the environment. I raise two problems for the argument. The first problem is that it is more difficult than Clark and Chalmers think to set up the Tetris example so that application of the parity principle might render it a case of extended mind. The second problem is that, even when appropriate versions of the argumentative examples can be constructed, the availability of a second, internalist parity principle precludes the possibility of inferring that the mind extends. Choosing which parity principle we ought to wield would involve deciding beforehand whether or not the mind can extend. Thus Clark and Chalmers beg the question by employing their parity principle rather than the internalist one. I conclude that they fail to provide a proper argument to support the extended mind thesis.

Introduction

I first explain Clark and Chalmers (henceforth: C&C)'s argument for the extended mind thesis (henceforth: EM), which is based upon their argumentative examples concerning Tetris and Otto, plus their 'parity principle'. Next I look closely at the Tetris example, and show that it requires some tidying up before it has the right structure to operate in an argument for EM. This tidying completed, I present a serious problem that afflicts C&C's argument, even with the best versions of the argumentative examples in hand; in short, C&C's use of the parity principle begs the question against internalism. I conclude that although C&C can be credited with raising EM as an interesting hypothesis, they have not provided a well-motivated (as in non-question begging) argument for its truth. Internalists will see no need to

take seriously an argument that builds the truth of externalism into its premises. What C&C do bequeath us however, we will eventually see, is an interesting dilemma concerning the scope of mentality.

The Argument for the Extended Mind

C&C frequently express EM by saying that a person's mental states supervene not just on goings-on within her skull, but on some goings-on in the environment as well. But supervenience is clearly not a tight enough relation with which to express EM, or to ground an argument for EM's characteristic claim, that the mind extends beyond the head. To see this, imagine that our mental states 'supervened upon' God's whim—that his state of mind was determinative of our own. Or, alternatively, consider a substance dualist who maintains that the state of his mind supervenes on the state of his body.ⁱ In neither of these cases is it desirable to include the supervenience base literally in the mind, and this prospect is made no more desirable just because the supervenience relation holds. So supervenience is by the by as far as EM is concerned. C&C must mean something much stronger, it follows, when they ascribe mentality to processes outside of the organism. What they really intend, of course—as they elsewhere explicitly state—is the *realisation* relation: the mind is held, on their view, to be partially realised in the environment.

As I construe it, C&C's argument that the mind is partially realised in the environment depends upon the application of their 'parity principle' to key examples, each of which features a pair of functionally isomorphic states or processes; one existing inside the organism, the other outside.ⁱⁱ

In their Tetris example, C&C compare the shape-rotation that a conventional Tetris player achieves manually on a videogame screen with imaginative visualisation of the same rotation.

In their Otto example, C&C compare the storage of endorsed content in a notebook with the storage of endorsed content in the cranial biological memory. What makes it plausible that the internal and external resources in each case are functionally isomorphic is the fact that the agent concerned achieves his end just as well and more or less just as easily via the use of either one. Tetris can be played just as well and more or less just as easily on a Gameboy screen as in the head, and Otto can locate the Museum of Modern Art in New York just as well and more or less just as easily by consulting a notebook as by consulting his onboard memory.

Given that the internal resources in the argumentative examples—mental rotation of shapes, and cranial storage of standing beliefs—are considered by C&C to be unproblematically mental phenomena, they utilise the apparent functional equivalence of these resources with their external analogues in order to stretch ascriptions of mentality out into the environment. They achieve this result through the application to the examples of what they call *the parity principle* (henceforth: PP), expressly intended, says Clark, to free us from our ‘biochauvinistic prejudice’^{iii,iv}—our unjustified tendency to ascribe mentality exclusively to *intra-organism* processes:

PP: ‘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/mental process, then that part of the world is (for that time) part of the cognitive/mental process.’^v

On this basis, rough functional equivalence between internal and external resources plus the parity principle, C&C conclude that manually rotating tiles on a videogame screen and storing accessible and endorsed content in a notebook are cognitive/mental processes; cases of the mind’s partial realisation in the environment.

Tidying up the Tetris Example

There is an important difference between mental rotation of Tetris shapes and the manual rotation one achieves on a videogame screen, a difference that threatens to wreck C&C's use of the Tetris example. Mental rotation of Tetris shapes is, notably, a *conscious* process: indeed, its intrinsic conscious self-presentation to the subject will be essential for conducting her imagination-based game of Tetris. Roughly speaking, this playing subject visually imagines the Tetris tiles rotating in order to identify them, assess their prospects for integration into the growing wall, and then (mentally) place them appropriately. When playing Tetris on a Gameboy on the other hand, the screen is not a resource of which the player is immediately aware; she is not directly conscious of this representation of the game. The representation of the Tetris game on the Gameboy screen is one thing, this player's conscious experience of the game another. If we take PP at its strict letter, therefore, the Tetris example gets nowhere near to establishing EM: for it is not the case that, were something like a Gameboy running in the Tetris player's head, this would by itself help constitute a cognitive/mental process. Simply having a videogame going on inside one's skull hardly counts as cognitive/mental activity!^{vi} The external process in this case is not such that were it occurring in the head we would count it as cognitive/mental. PP does not, therefore, recommend counting a Gameboy on which one is playing as a cognitive/mental resource.

In their two cases intermediary between uncontroversially *mental* (imaginative) shape-rotation and conventional (external) onscreen Tetris play, C&C first envisage a Martian with a biologically inherited motor-controlled retinal display, and then a cybernetically engineered near-future human with a similar retinal display who can choose between using this or employing mental visualisation to rotate Tetris shapes. C&C claim that the Martian's inherited apparatus is indisputably a cognitive/mental resource, and that this encourages the

same classification of the technological endowment of the near-future human. By functional parity—via PP—therefore, the shape-rotation a conventional Tetris player accomplishes on an external Gameboy screen also counts as cognitive/mental activity, activity realised in the extra-cranial processes of the Gameboy. In light of the difficulty we identified above for C&C’s claim that conventional Tetris play can constitute an instance of EM, it is worth lingering over the two intermediate cases to see whether we might glean from them a clearer idea of how to read the Tetris argumentative example in its strongest possible form.

Considering the intermediate cases further, we can see that there is some awkwardness in the fact that the internal display-screens are located at the retina. For it is not obvious that the retina counts as a cognitive/mental resource; when one has a scratched retina, for instance, it does not seem to follow that one has a scratched cognitive system/mind. Of course, *if* the mind can extend then perhaps the retina could become part of it, under the right circumstances. But that possibility cannot be assumed here without begging the question, for we are still in the stages of the argument for EM. This means trouble when we apply PP: were some environmental process to operate in a manner functionally isomorphic with the hypothesised retinal displays, this would not encourage us to think of that process as cognitive/mental, on account of the fact that its internal analogue is not cognitive/mental to start with. The external process would not be such that were something like it going on in the head we would count it as cognitive/mental. Hypothetical retinal displays are therefore a bad place to start if one wishes to build a C&C-style argument for EM.

This difficulty is easily overcome however, we need only re-cast C&C’s retinal display as a display-screen realised somewhere further inside the brain, in an uncontroversially cognitive/mental system; within the visual cortex for example. Having taken the screen away from the eye, though, we now face clearly the crucial question about the subject’s awareness

of what it represents: are we to imagine the screen set up so that (somehow) the subject is directly conscious of what is on it, so that the screen in some way *constitutes* the player's conscious awareness of the Tetris game? Or are we to imagine things so that the screen merely *causes* (not constitutes) the player's conscious visual representation of the game? (For the latter interpretation, picture a Gameboy screen embedded in the player's skull and then suitably patched into her visual system,^{vii} so that she experiences just the same visual imagery of the Tetris game that she does when holding a Gameboy close to her face)

The constitution reading creates a new problem for C&C's application of PP in support of EM—for it opens up a new discrepancy between the internal and external resources: It is not the case that were a Gameboy implanted (however ingeniously) into someone's head it would constitute their conscious representation of a Tetris game. On the constitutional reading, therefore, even if the Martian/near-future human's internal display is a cognitive/mental resource, PP does not recommend counting a normal (external) Gameboy as cognitive/mental also: for the latter is not such that were it running in the head we would consider that a cognitive/mental process. Such an internal process would need to be a conscious representation of the Tetris game, but a Gameboy screen is not itself a conscious representation of anything, wherever it is.

C&C must—or at least ought to—intend the causal reading of the Martian/near-future human cases, it follows: we are to think of the internal displays as causing these Tetris players' conscious representations of how things are with their games. And now we seem to have achieved the desired internal/external functional isomorphism: for a normal Gameboy in play is also the cause of the player's conscious representation of the game she is playing. We have succeeded in tidying up the Tetris argumentative example: As long as the internal cause of a playing subject's conscious representation of her Tetris game counts as a cognitive/mental

resource, PP will encourage us to count an external cause of such a conscious representation—something like a Gameboy—as a cognitive/mental resource as well. In this way EM will be supported.

However, a serious objection faces this, the optimum version of the Tetris example, an objection that applies equally to the Otto example, to which we turn next.

Otto, Tetris and Parity Principles

The difficulty for C&C's use of the Otto example is that so-called 'parity principles' are, it turns out, relatively cheap, and easy to come by. Here is a second one:

PP* 'If, as we confront some task, a part of the head functions as a process which, were it to go on in the world, we would have no hesitation in rejecting as part of the cognitive/mental process, then that part of the head is (for that time) not part of the cognitive/mental process.'

This principle yields the conclusion that standing beliefs, stored in the organism, are not part of the subject's cognitive/mental process. For we do not *prima facie* regard content-storing states of the environment, whether accessible to consciousness or not, as mental. On the other hand, consciousness is clearly not disqualified as mental by PP*. Though we may consider it naturally impossible, there is little conceptual trouble in making sense of a subject's consciousness partially, or even fully, extending outside of her organism. This is why it is not *a priori* obvious that 'out of body' experiences are illusory. In such a case we would accept the verdict that the subject's mental process had extended: thus consciousness passes muster according to PP*. And PP* also makes the correct adjudication about processes such as the growth of myelin sheaths (i.e., that they are not mental). Application of this principle to Otto does not establish EM; PP* would instead, applied broadly, help establish the thesis that Gertler^{viii} calls 'the narrow mind', the view that it is only occurrent consciousness that is

properly-speaking mental. For it is, arguably, only where we find consciousness that we are compelled to posit mentality.^{ix} Otto would, on this reading, only have the content that the MoMA is on 53rd street ‘in mind’ when he reads his notebook and consciously entertains it.

Returning now to the tidied Tetris example: the application of PP* here would also support an internalist conclusion, rather than EM. For it is not *prima facie* the case that an external screen that causes a conscious representation of a Tetris game in the player’s mind—something like a Gameboy—counts as a cognitive/mental resource. By PP*, therefore, an internal analogue of such a screen would not count as cognitive/mental either. For it would be a process such that were it going on in the environment we would not count it as cognitive/mental. Let me be clear: this is not to deny, far from it, that internal *conscious* visualisation of a Tetris game is a cognitive/mental process; indeed I think most would agree that such a process is paradigmatically cognitive/mental. What is denied cognitive/mental status by PP* is, rather, any internal process merely *causally underlying* this kind of conscious visualisation. For, to repeat, external causes of conscious representations are not themselves taken to be cognitive/mental phenomena.^x On this reading of the Tetris example, neither the near-future human nor the Martian possess genuinely cognitive/mental resources in the form of their internal Tetris displays.^{xi}

What is revealed by the availability of PP* is a disingenuous element in C&C’s treatment: the parity principle they employ is not as even-handed as its name suggests; it is far from a “‘veil of ignorance” style test^{xii} provided only to help us neutralise our organismic bias, as Clark suggests. Certainly PP might appear quite harmless, because at a glance it seems to amount to no more than pointing out an equivalence between two types of item: an extra-cranial resource and an intra-cranial one. And equivalences do not have priorities or biases built-in of themselves, we normally think. But it is also clear that, depending upon the interests of the

person drawing a particular equivalence, it *is* possible to build priority into that equation; this hangs on which side of the equation the person takes as basic. The equation can thereby be imbued with directionality; bias. It is an analogous point that is at the root of anti-physicalist complaints that physicalist reductionists are in effect proposing a kind of eliminativism about the mind, for example. Reductionists claim to be proposing no more than an equivalence when they say that mind ‘just is’ matter. But anti-physicalists find that when the material side of the equation is taken as basic, the explanandum that matters to them can seem to disappear, as if the victim of a conjuror’s trick.

C&C are also engaging in conjuror’s trickery with their use of PP. PP is ingrained with externalist bias; its application to the examples is designed to yield EM as conclusion. I do not claim that PP* is any more innocent (though note that it can be given its own egalitarian-sounding gloss, as intended to help us avoid another kind of biochauvinistic prejudice: the tendency to think that if some resource is skull-based then it is *ipso facto* cognitive/mental; PP*, as much as PP, encourages us to ignore the boundary of skull and skin when it comes to attributing cognition/mentality, as Clark so earnestly recommends). Given its relation to PP, PP* wears its internalism on its sleeve; it is designed to reverse the verdict of EM and radically restrict the scope of mind.

PP and PP*, then, cancel one another out dialectically. We might at this point, I think, usefully introduce a third, second-order parity principle:

PP² ‘If, as we confront conflicting first-order parity principles, each principle can equally be given an egalitarian gloss, and each delivers equally on substantive but unobvious conclusions, then we should have nothing to choose between them *prima facie*.’

In PP², I submit, we finally have a parity principle that operates without prejudice. And what it indicates is that given a choice of inconsistent first order parity principles, we need some independent reason to endorse one rather than the other, in the sense of selecting it for application to argumentative examples of the sort we have been considering, and which C&C take to support EM. The question we now face, consequently, is this: Why should we choose to employ PP instead of PP*, thereby arriving at C&C's externalism, rather than proceeding vice-versa, a path that will take us towards an equally radical internalism?

Does PP² provide us the means to decide between PP and PP*? I do not think so: by the criteria of PP² the two first-order parity principles are on a par. The conclusions they each support, that the mind can be partially realised in notebooks, Gameboys and other environmental resources, and that the mind is limited to occurrent, conscious content (so that standing beliefs are non-mental) are equally surprising and radical, and from an impartial perspective each thesis departs equally far from the common sense view of the mind (which I daresay has it both that the mind cannot extend and that standing beliefs are mental). If we are to resolve the issue of which parity principle to wield, then, we will have to look for our reasons elsewhere.

Why the Argument for the Extended Mind Begs the Question

If we apply PP to C&C's argumentative examples we arrive at EM. If on the other hand we apply PP*, standing beliefs turn out not to qualify as mental, and although imaginatively visualising a Tetris game is undoubtedly a cognitive/mental activity, no relevantly analogous—that is, conscious—activity can occur outside of a head, as C&C themselves concede.^{xiii} PP* urges upon us a much stricter conception of what qualifies as genuine mentality.

We need some means of deciding which (first-order) parity principle we ought to apply to the argumentative examples, in order to know which way the argument for EM will go.

But, I submit, there is simply no way to do this without sorting out in advance whether we think EM might be true or not. Those who will accede to applying PP to C&C's argumentative examples will be those antecedently sympathetic to the idea that mental processes might extend. Conversely, those with strong internalist leanings will naturally want to apply PP* to the examples. Further, if we think for a moment about the kinds of consideration that we might adduce to encourage someone to accept that it is (say) PP rather than PP* that should be applied to a given example, we see that these will precisely be considerations that already stand in favour of EM. And conversely. One will only find such considerations persuasive, it follows, to the extent that one is already disposed towards or against EM.

It seems that to decide which parity principle properly applied to a given argumentative example (consisting of a pair of functionally isomorphic putatively cognitive/mental resources, one internal one external), one would have to have already made up one's mind about EM's truth. It follows that PP cannot be used in an argument for EM; rather, it could only be the outcome of such an argument. And since C&C rely on the employment of PP to drive through their externalist conclusion, rather than a principle like PP*, we must conclude that their argument for EM begs the question.

Conclusion

C&C have certainly raised an interesting possibility, that the mind extends; but they have offered no serious (as in, non-question begging) argument in favour of the thesis. The real argument over EM, it seems, is yet to come, for both sides in the debate. We might end by

noting a further interesting consequence of C&C's argument, given the options we now have on the table: it appears that we are forced either towards a radically externalist or a radically internalist conception of genuine mentality. Even if the correct choice (in so far as our concept of mentality admits of one) remains as yet undetermined, having uncovered this surprising dilemma would represent significant progress in the philosophy of mind, and for that C&C deserve credit.

ⁱ David Chalmers considers such a position, for example, under the rubric 'naturalistic dualism'. See his *The Conscious Mind* (Oxford University Press, New York, 1996).

ⁱⁱ The argument is first proposed in Andy Clark and David Chalmers "The Extended Mind" *Analysis* 51:8:7-19 (1998). See also Andy Clark *Supersizing the Mind* (Oxford University Press, New York, 2008) for the latest formulation and defence of EM, which clarifies the original argument structure. It is Clark's reworking that provides my main target here.

ⁱⁱⁱ Clark (op cit) p77.

^{iv} In fact, in the original paper PP is introduced on the basis of *first* considering the Tetris example. Nonetheless, C&C make it clear that it is PP that (also) drives their reasoning about the Tetris case.

^v Slightly adapted from Clark and Chalmers (op cit) p8. In the original PP the term 'mental' does not figure: C&C took the Otto example to establish extended mentality (because of the assumed mentality of internal standing beliefs), and the Tetris example to establish EM to the extent that cognition implies mentality, for they take this example to establish extended cognition at least. I set aside this difference between the examples here, as well as the question of any shortfall between the cognitive and the properly mental; nothing substantive hangs on these issues for my present purposes.

^{vi} Of the right kind anyway; it would surely have all sorts of unfortunate cognitive/mental effects

^{vii} Plus motor-cortex, as before, so that she can control the game.

^{viii} Brie Gertler "Overextending the Mind?" in B. Gertler and L. Shapiro (eds) *Arguing about the Mind* (Routledge, New York, 2007).

^{ix} Gertler is far from alone in this controversial view of mentality; Strawson, for example, is another important advocate (see Galen Strawson *Selves: An Essay in Revisionary Metaphysics*, (Clarendon Press, Oxford, UK, 2009)). I take no stand on the matter here; my key point in this paper is simply that nothing C&C say prejudices the issue against this kind of internalism.

^x Except where we have independent reasons to consider them cognitive/mental; e.g. when they are other people

^{xi} When these are taken according to the causal reading of the scenarios. Recall that while on the constitution reading the internal displays might count as cognitive/mental, they make functional isomorphism with an external Gameboy—and so, the application of PP—impossible.

^{xii} Clark (op cit) p77.

^{xii} See for example Chalmers's foreword to Clark (op cit).