

Extensive Enactivism: Why Keep it All in?

Daniel D. Hutto, Michael Kirchhoff and Erik Myin

There are people of intelligence who can *learn* as many of
the facts of science as they like, but
... they lack the spirit of science.

For them it is enough to have discovered any hypothesis at all concerning any matter,
then they are at once on fire for it and believe the whole thing is accomplished.

To possess an opinion is to them the same thing as to become a fanatical adherent to it,
and henceforth to lay it to their heart as a conviction ...

Insofar as genius of every kind maintains the fire of convictions
and awakens distrust of the modesty
and circumspection of science, it is an enemy of truth,
no matter how much it may believe itself to be truth's suitor.

- Nietzsche, *Human all too Human* I: 635

1. Getting Beyond Extended Minds

After years of debate, philosophers and cognitive scientists are still divided over the question of the extent of cognition. Does cognition happen entirely in the brain or is it, instead, a dynamic and interactive phenomenon that constitutively involves the environment, both spatially and temporally? Clark and Chalmers (1998) created recent interest in this discussion in their seminal paper “The Extended Mind”. Despite the energetic production of publications arguing back and forth over the issue since then, it has been proposed that the extended mind debate is in a deadlock (Rowlands 2010, p. 210). Neither side appears to have decisive means of settling the matter in its favour.

In *Radicalizing Enactivism*, Hutto and Myin (2013) articulate and defend a radically enactive and embodied – thoroughly non-contentful – vision of basic cognition (REC for short).¹ The book defends the view that accepting REC has the advantage of enabling us to transform the terms of the extended mind dispute in the philosophy of mind. Specifically, its seventh chapter argues that endorsing REC provides both a spur and the sufficient means of moving beyond the ‘extended mind’ debate, and its alleged stalemate. The book’s diagnosis is that a fundamental obstacle to progress in the extended

mind dispute is a continued commitment to the idea that cognition is necessarily content involving (CIC for short), something that has been accepted by those on both sides of it (Hutto and Myin 2013, pp. 135-136). According to a REC analysis, if we let go of the CIC assumption our understanding of the extent of cognition is transformed, requiring a complete reconceiving of the standard extended mind debate. If we are correct, giving up on the idea that minds are essentially representational – always and everywhere content involving – is a game changer.

The reasoning is straightforward. Assume that basic cognition is not representational in a contentful sense. If so, then the clearest and cleanest grounds for internalism are undercut. For if basic cognition is not inherently contentful then defenders of internalism lose their most compelling reason for supposing that cognition – at least in its primary phylogenetic and developmental forms – is an ‘inner’ business (where ‘inner’ is understood in the sense of occupying a cerebral location). These observations about how the extended mind debate would be transformed in REC’s wake seize on the fact that the standard, and strongest, move internalists can make to motivate their position is to appeal to a notion of narrow or intrinsic content (Adams and Aizawa 2010). The appeal to mental content features crucial in internalist arguments because it is needed to provide a principled ‘mark of the cognitive’ – one which backs up and gives definition to demarcation claims about what is constitutive of, as opposed to merely causally supportive of, cognition. This assumption is widespread and easy to find in the literature. The following quotations epitomize familiar sentiments about the representational mark of the cognitive and its perceived importance:

Admittedly, delimiting the scope of the ‘cognitive’ is not an easy matter, but ... it seems adequate to specify that cognitive states, structures, and capacities are mental entities with representational content (Khalidi 2007 p. 93).

Without representation cognitive science is utterly bereft of tools for explaining natural intelligence. We would go further: without representation there is no cognitive (as distinct from behavioral, biologic, or just plain physical) science in the first place (O’Brien and Opie 2009, p. 54).

Anyone who wants to make claims about the extent of cognition and abandons the representationalist ‘mark of the cognitive’ must supply an alternative, otherwise there is no way of distinguishing the cognitive from the non-cognitive.

Ultimately, we agree with O’Brien and Opie (2009). In the absence of an appeal to content there is no obvious alternative way to ground claims about what constitutes the cognitive in a scrupulous scientific manner. As we aim to show, one consequence of this is that without appeal to a notion of content to supply the mark of the cognitive – or an adequate replacement that can play that role – there is no principled way to advance the claim that cognition is neurally housed. How else, other than

by appeal to content, might the claim that cognitive processes are ‘contained in the brain’ be supported? How could it be established that, as a matter of fact, cognition is always and necessarily brainbound? Appeals to facts about brains or behaviour on their own, without supplement by a substantive theory of cognition, would not warrant that conclusion. We will return to this.

From a REC perspective, to let go of the idea that basic cognition is necessarily contentful and representational in character is to remove a standardly assumed barrier to seeing cognition as constitutively world-involving. Such a shift in perspective is tantamount to acknowledging that fundamental cases of perceiving and thinking are not necessarily grounded in, nor do they take the form of, representing aspects of the world or having contentful thoughts about those aspects. And to think this aligns perfectly with understanding cognitive processes as a form of wide reaching activity that is – at root – extensive and unbounded; thus extensive minds are not merely, occasionally and in special circumstances, extended.

2. Insufficiency and Irrelevance

Not everyone agrees that going radically enactive about cognition has these negative implications for internalism or these positive implications about the wide extent of cognition. In this section we focus on two charges that have been leveled, not against the truth of REC, but its significance in this domain.

In a closing keynote to the *Varieties of Enactivism: A Conceptual Geography* symposium, hosted by the AISB-50 in April 2014, Wheeler claims that enactivism, even its non-representational variants, lacks the requisite theoretical resources to see off internalism about the machinery of cognition (or as he describes it the ‘whereabouts of our cognitive architecture’). Wheeler maintains that REC’s claims about the extensive and constitutively world-involving relationality are insufficient for rejecting internalism: by his lights, when it comes to putting internalism out of business REC “falls short of what is needed” (2014, p. 1).

He holds, by contrast, that extended functionalism – all on its own and without the aid of non-representationalism – suffices for rejecting internalism: extended functionalism, if he is right, does the trick whether it endorses representationalism or not. Consequently, he claims that extended functionalism can carry the day in the extended mind debate while remaining studiously neutral or agonistic about the question of the representational nature of cognition. In taking this line Wheeler denies the strong connection that RECers see as holding between representationalism and the rejection of internalism.²

Sutton (2014) also doubts that non-representationalism is relevant to debates over the extent of cognition. He holds that whether internal representations are assumed to exist or not is orthogonal to the internalism-externalism dispute – at least, if one’s unit of analysis is cognitive activity understood as a public, distributed process. If Sutton is right, explanations of distributed cognitive activity (his

examples are taken from the science of memory) can proceed without our having to get clear about, let alone settle questions about the existence (or otherwise) of mental representations. Although he doesn't always explicitly pitch his detailed discussions in terms of philosophical debates about mental representations, for him it remains an open question to what extent, for example, a "radically revised notion of the internal memory trace" commits theorists to mental representationalism (Sutton 2014, p. 5). In moving memory research ahead he denies that we need be delayed by debates of this kind, for that is not where the real action lies. Hence Sutton's verdict is that debates about mental representations between RECs and their opponents are orthogonal when it comes to moving forward in (a) the internalism-externalism debate and (b) the explanatory project of understanding cognitive activity as a socio-culturally distributed process.

We disagree with both Wheeler (2014) and Sutton (2014). But before scrutinizing their views, let's consider one diagnosis of why non-representationalism may be thought to be irrelevant to the extended mind dispute. It might be thought that this follows because *all* issues concerning mental representation are irrelevant to that debate. Why think so? Here's a line of reasoning. Externalism and internalism take different forms depending on one's theory of mental representations. Different varieties of representational theory assume that representations, whatever specific properties they happen to have, must – as a class — have *content* (of some kind) and *vehicles* (of some kind) if they are to qualify as representations at all. The qualification 'of some kind' is important for it reminds us that it is sufficient only that the vehicle-content distinction is in play: it does not matter which particular form it takes. Against this backdrop it is possible to capture the main theoretical combinations in the internalism-externalism debate using a fourfold matrix. One can pair: (i) vehicle-internalism with content-internalism (Fodor 1990); (ii) vehicle internalism with content-externalism (Fodor 1994, Dretske 1995); and (iii) vehicle-externalism with content-externalism (Wilson 2004). It is also logically possible to pair (iv) vehicle-externalism with content-internalism – though we know of no one who adheres to this view.

What does this set of possible pairings show? It might be thought to reveal that questions about mental representations are orthogonal or irrelevant to what one thinks about the extent of cognition precisely because in adopting any one of the four combinations about representational vehicles and content one will occupy a position on either side of debate. One's choice in this regard decides where one stands in the debate (though options (ii) and (iv) raise interesting questions), but that choice, on its own, cannot settle the outcome of the debate. So it might be thought that one must look to factors above and beyond representational factors if one is to move things along.

But to conclude from this line of reasoning that representationalism is irrelevant to the extended debate would be a mistake. Why? Firstly, if decisive arguments could be found to favour one of these four theories of representation over its rivals it would be possible to bring the discussion to a close. So, far from being irrelevant, considerations about mental representations look as if they lie at the very heart of this dispute. Secondly, and more fundamentally, if representationalism is rejected *tout*

court – if the issue of whether cognition is representational *at all* is moot – then the common ground for the extended mind debate, as characterized by this matrix, collapses under the very feet of all parties.

Consider an analogy. Several political parties disagree about the appropriateness of introducing new tax laws – each canvassing different combinatorial options. None of these parties question the need for taxation. Of course, merely recognizing that there is a need for taxation would not decide which combined tax law should be adopted. Other considerations would have to be brought to bear. Questions of taxation would not be irrelevant if those other considerations turned out to be internal to the choice between the various tax options on offer. Foundationally, the agreed need for taxation could not fail to be relevant to the imagined political debate, for it is the very basis upon which it is conducted. And, obviously, were taxation to be abolished the entire debate about different possible tax laws would be pointless. By the same token, it is difficult to see how one's stand on mental representation could be irrelevant or orthogonal to the extended mind debate.

3. From Extended Functionalism to Extensive Enactivism

Having neutralized one reason for thinking that mental representations are, after all, relevant to the extended mind debate, we need to consider exactly what it might mean to adopt a non-representationalist extensive enactivism and what might motivate doing so. First things first. In Hutto and Myin 2013 it was argued that understanding cognition as extensive and not extended is a consequence of adopting REC. Questions have been raised about the very idea of extensive minds. Wheeler (2014) wonders: "What does this mean?" (p. 1). By way of reply it is important to note that coining the label 'extensive' was meant as a corrective – a twist designed to highlight both what was, at once, right and wrong with the 'extended mind' metaphor. How are these notions different? Wheeler captures it perfectly, when he asks: "is mind a phenomenon whose primary ontological manifestation is inner, but which sometimes spreads beyond the skin (extended functionalism as often understood), or a phenomenon that is widely constituted in its very essence (the extensive mind of radical enactivism)[?]" (p. 2).

As a first pass answer, we think this reply makes the difference clear enough. Yet there are some who, despite admitting that the original 'extended mind' terminology has its difficulties, resist taking up our neologism. Sutton (2014), for example, thinks following our lead on this score is unnecessary. As he sees it, other serviceable replacements are, and have long been, readily available in the literature.

It is true that the word 'extended' can easily be misread as assuming a more basic inner cognitive system which only spreads later in development. For this reason, the well-established pre-existing label 'distributed cognition' should be preferred (Hutchins 2014): there is no need for the awkward enactivist coinage 'extensive mind' (Sutton 2014, p. 11).

Sutton's suggested swap of labels runs into trouble as long as we conceive of the disagreement between internalists and externalists as a metaphysical debate about the extent and boundaries of the cognitive. For in that case the notions of extended mind and distributed cognition will be mismatched. This is because distributed cognition – as Hutchins defines, and Sutton endorses, it – denotes a pragmatic stance not a metaphysical claim. Thus “to take the distributive perspective is *not to make any claim about the nature of the world*. Rather, it is to choose *a way of looking at the world*, one that selects scales of investigations such that wholes are seen as emergent from interactions among their parts” (Hutchins 2014, p. 36, emphases added).

As a pragmatic stance distributed cognition may well be the right explanatory attitude to adopt. It is surely compatible with extensive enactivism. But it is also equally compatible with internalism. How so? This follows trivially if distributed cognition makes and defends no claims about the nature of cognition – for if that is right then it does nothing to oppose internalism. Here's the rub: if distributed cognition makes no claims about the nature of the world then it has no stake, or justified say, in the extended mind debate construed as a metaphysical disagreement about the boundaries of cognition. Fans of distributed cognition can legitimately avoid entering into the discussion if they can provide compelling reasons for thinking that metaphysical disputes in general do not matter or that the extended mind debate is not, at root, a metaphysical issue. There is room for disagreement on this score (see Hurley 2010). But for anyone who regards the extended mind debate as a metaphysical debate the labels ‘extended mind’ and ‘distributed cognition’ are not interchangeable and speaking in terms of distributed cognition does nothing legitimate to silence the internalists.

From a metaphysical perspective, we believe the notion of extensive mind can and should replace talk of extended minds. But to explicate properly we need to go beyond the first pass characterization, for some still find the notion of extensiveness imperspicuous. There is a need, as Wheeler (2014) identifies, to say more about “precisely how are we to explicate the property of extensiveness, and thus what it means for mind to be “widely constituted in its very essence”?” (p. 2).

By way of a more developed answer let's look at how extensive enactivism can be derived from extended functionalism and how this follows if one surrenders the idea that basic cognition involves contentful representation. What's the link? Roughly rendered: Non-representational functionalism, extended or otherwise, is a pure form of functionalism. Pure functionalism does not entail any kind of representationalism: pure functionalism is a weaker thesis than representational functionalism.

We will now show that there are reasons to think that if extended functionalism is advanced as a variety of pure non-representational functionalism it transforms into extensive enactivism. In what follows we will trace the steps of how this process unfolds. Through this analysis it will be shown how and why extensive enactivism can be understood as extended functionalism transfigured. Of course, if things go this way, there will be no question of favouring extended functionalism over extensive enactivism. Capitalizing on Wheeler's penchant for Potterisms, it is instructive to observe,

as Hermione Granger reminds us, that “*Transfiguration, you know, turning something into something else, of course, it’s supposed to be very difficult*”(Rowling 1997, p. 93-94).³ In this case, we don’t think the process is all that difficult or magical. It can be transparently achieved by a bit of philosophical analysis.

In reviewing the fate of non-representational, pure extended functionalism we will thereby explicate the notion of extensiveness and it will become clear that asking questions about boundaries of the cognitive is always a matter of asking questions that take stock of the world-relatedness of cognition.

4. Extended Functionalism via Empirical Functionalism

Without the backing of representationalism pure functionalism encounters intractable problems of definition and demarcation – it loses its principled basis for determining what cognitive ‘inputs’ and ‘outputs’ are, and thus for saying where cognitive processes might begin and end. Without a substantive theory of cognition to provide requisite backing it is not possible to determine the significance that various facts about brains, behaviour or computation have for our thinking about the extent of cognition. Abandoning the idea that cognition is representational – content involving – leads step-by-step to an unthreading of empirical functionalism and ultimately brings it face-to-face with old and familiar charges of its triviality.

There is a familiar story about computers according to which computational operations are performed on various bits of information stored as unique strings of code. These bits of information are conventionally identified by a unique assignment of 0’s and 1’s and at the hardware level these ‘codes’ are said to be physically realized in the computer’s on-off states. This ensures that a computer’s operations can be carried out purely mechanically. It also explains the great versatility of computers, which derives from the fact that these various sets of 0’s and 1’s can be given different semantic interpretations. In one computer programme what is found at a given register might be treated as a word, in another it might be treated as a number, or in yet another, as some geometric figure. That’s the upside. The downside is that “If computation is defined in terms of the assignment of syntax then everything would be a digital computer, because any object whatever could have syntactical ascriptions made to it. You could describe anything in terms of 0’s and 1’s” (Searle 1990, p. 26). Call this the problem of computational individuation.

A way around this problem is to make good on a theory of computation of the sort offered by Piccinini (2008) – one that understands computation in purely nonrepresentational, functional terms. This requires giving a theory or explanation that is “sufficient to individuate computational states without appealing to either semantic or syntactic properties” (Piccinini 2008, p. 209).

We agree that a theory of this kind is what is needed. But even if such a theory provides a basis for a pure functionalist account of computation there is a need to forge a further link so as to connect that

purely functional account of computation to cognition. In assessing the prospects of discovering that link it is crucial to realize just how wide and unbridgeable the gap in question really is if one adopts a non-representational notion of computation.

We can get a grip on the magnitude of the problem by seeing how different and distant a purely functional theory of computational individuation is from either a representational or syntactic theory of mind. Consider that the received view takes it for granted that, “computational states are individuated at least in part by their semantic properties” (Piccinini 2008, p. 205). Or, as Fodor more pithily puts it: “there is no computation without representation” (Fodor 1981, p. 180). Or, again, as O’Brien and Opie (2009) maintain “computation is governed by the contents of the representations” (2009, p. 53).

When it comes to individuating mental states orthodox cognitive science has lived by the hope that the threat of triviality generated by the problem of computational individuation could be avoided by a different route than that proposed by Piccinini (2008). The problem might also be dealt with if the hypothesized symbol-strings had *bona fide* representational properties. If that were so classical cognitive scientists would have a toehold for claiming that a theory of cognition that postulates symbols, with both semantic and syntactical properties, might be true of organic brains.

If a unified computational and representational theory of mind could be developed to provide a justification for the realistic assignment of contents it could thereby individuate syntactic structures. This proposed link is clearest in the language of thought hypothesis that invites us to imagine, put crudely, “that human mental sentences are written on little cerebral CRTs” (Stich 1983, p. 53). Yet what if we abandon the idea that the Mentalese sentences that are imagined to fill our heads are contentful? Can we just drop the idea that there is a relation between representational semantics and computational syntax by rejecting the former while hanging on to the latter?

In order to avoid familiar worries about the casual impotency of content Stich (1983) proposed doing just that. He advanced a syntactical theory of mind according to which no such relation existed; that is to say, he proposed a retreat to a pure, non-representational functionalism. But to abandon the idea that mental states have semantic properties raises questions about the very notion of syntax. What does a proponent of this view imagine is left behind? Syntax and semantics, as classically understood, are internally related: thus, it is simply not conceivable that talk of syntax in the absence of semantics, let alone how we might begin to individuate mental states and carve up the mind in purely syntactical terms without appeal to semantic properties. Thinking that this separation is possible is a “common but serious mistake” (Piccinini 2008, p. 208).⁴ It will only appear unproblematic that we can talk of syntactic properties entirely independently of semantic properties if we imagine that syntax keeps its shape, somehow, as a kind of shadow of semantics. Yet once semantics goes, its shadow – syntax – goes too.

We are thrown back on the old problem that pure functionalism always faces: How are we to define and individuate mental states without appeal to representational properties? Certainly, we

cannot do this by appealing to bottom-up facts of neuroscience in the absence of a substantive theory of cognition. Even Adams and Aizawa (2010) acknowledge that:

We have no way to identify particular tokens of brain states qua syntactic state items in order to affix contents to them. Given the state of current science, we only identify a person's brain states via inferences to the content of those states (p. 72).

Our analysis takes this observation a step further: we cannot identify a person's brain states qua syntactic states except by appealing to semantic properties. Once this is acknowledged it looks as if a retreat to pure functionalism leaves us bereft of the resources for individuating the cognitive. As the infanticidal father of functionalism writes of his once brainchild:

One looks for something definable in nonintentional terms, something isolable by scientific procedures, something one can build a model of ... And this – the “mental process” – is just what does not exist (Putnam 1988, p. 74).

Is this correct? Let's address this question by examining some recent proposals in the extended mind debate. Adams and Aizawa (2010) advance an internalist proposal about how to individuate the cognitive by appeal to two necessary conditions: (i) the intrinsic content condition and (ii) the causal processing condition. Accordingly “cognition is constituted by certain sorts of causal processes that involve non-derived content” (p. 68, see also Adams and Aizawa 2001, p. 52-53). Of course (i) has to go if we give up on representationalism. What's left? Could appeal to (ii) on its own provide secure basis for thinking that cognition is a constitutively heady affair without an appeal to content? Without CIC in play, are Adams and Aizawa right to think that the ‘weight of empirical evidence’ favours cognitive intracranialism (p. 74)?

It is hard to see how this could possibly be so. For given how Adams and Aizawa set things up once we lose the first condition we are left only with the unilluminating thought that certain sorts of causal processes constitute cognition. That would be a kind of internalist pure functionalism. Perhaps this internalist functionalist proposal might be defended by appeal to the idea that we can demarcate which causal processes are the ones that matter by appeal to information processing differences. Yet, again, things get tricky if we give up on the idea that informational content is processed. Talk of information processing becomes less than perspicuous if no content is literally processed – once that favourite metaphor of cognitive science gives way (Hutto and Myin 2013, ch. 4). Indeed, if information is understood in purely covariance terms then an organism's sensitivity and responsiveness to information will be a world-involving activity, as extensive enactivism would have it. What would warrant dividing up such a ‘process’ into smaller cognitive and non-cognitive parts? How, without an appeal to content, could this lend any support for functionalist internalism? Adams

and Aizawa (2010) recognize this problem themselves holding that unless an appeal to content is in play the idea “that information processing constitutes the mark of the cognitive ... is implausible” (p. 76).

But perhaps it might be thought that all is not lost for internalist functionalism. What if there was a special way that brains respond to information intracranially – a way that, say, marks out such neural activity as being importantly different, setting it apart from the way brains behave when they are engaged in activities involving, e.g., extra-cranial brain-tool combinations? Rupert has developed arguments along these lines to support the claim that human cognition is located “inside the organism either entirely or in the main” (p. 45). To reach this conclusion he invokes the systems-based principle – a principle that makes no reference to representational content and thus suits pure functionalism. Accordingly:

[A] state is a cognitive state if and only if it consists in, or is realized by, the activation of one or more mechanisms that are elements of the integrated set [of members] which contribute causally and distinctively to the production of cognitive phenomena (2009, p. 42).

How does this help? There is a continuing, well known debate about the exact nature of mechanisms (Craver 2007, Bechtel 2008, Carver and Darden 2013). Let us imagine the best case for Rupert and suppose that all goes well with recent work on mechanisms and that it yields a precise means of individuating them from the rest of nature. Let us also imagine that, on that basis, we can determine precisely when, where and which mechanisms are implicated in cognitive activity. Assume further that we are reliably able to discern which are the more, or even the most, integrated mechanistic parts of any given cognitive activity. Even in this perfect scenario it is not clear what would license treating the non-mechanistic or less integrated parts of the activity as non-cognitive. That verdict would be driven by appeal to the stipulated systems-based principle. But lacking a substantive theory of cognition – as opposed to a theory of mechanism – what independently justifies that conclusion? For this sort of proposal to succeed the connection between the mechanistic and the cognitive needs be clearly established.

We can see the problem clearly by considering the fact that no one in the debates about the extent of cognition denies that cognitive activity involves mechanisms. What is questioned is why anyone ought to suppose that the cognitive is limited or restricted to certain, say, mechanistic or computational parts of such activity? Consider the analogy between perceiving and driving. Like driving, perceiving can be understood as a situated, environment involving activity – and nothing short of that. To be sure, driving depends, in part, on the activity and interactions between a set of integrated mechanisms. Moreover mechanisms are surely found within the car, but it does not follow from that fact that driving happens within the car. Nor is anything gained for mechanists by noting the fact that “one can manipulate the car’s behaviour by manipulating its engine” (Noë 2004, p. 211). For

it is, of course, equally true that one can intervene on the driving activity by ‘wiggling’ the (presumably) non-mechanistic environmental features, such as the condition of the roads.

Without some further grounds for conferring cognitive status on integrated mechanisms, other than stipulation, we face essentially the same problem again: nothing justifies the assumption that mechanisms *per se* pick out the cognitive.⁵

The root issue is that even if mechanisms feature in cognitive activity this fact does not, by itself, tell us what significance such mechanistic activity has for demarcating the cognitive from the non-cognitive. To see this, note that what is intended to justify the systems-based principle is *not* a substantive theory of cognition; rather, it is the alleged empirical fact that it is by presupposing that mechanisms are where the *real* cognitive action resides is what best “accounts for the successful practice in cognitive psychology” (Rupert 2009, p. 43). In fact, this defense of the systems-based approach rests on a number of disputable empirical claims. The most important are that: (1) orthodox cognitive psychology has been successful; (2) orthodox cognitive psychology individuates cognitive activity in terms of mechanisms; (3) that interesting laws of cognitive psychology are tied to finding mechanisms; and (4) the latter facts (2) and (3) conspire to best explain (1).⁶

Let us suppose, for the sake of argument, that (1)-(4), as stated above, are all true. That supposition poses no threat to thinking that cognitive activity is extensive. For it may *also* be true that cognitive psychology is or will be successful because it individuates cognitive activity non-mechanistically and that it will discover other interesting laws that are not tied solely to mechanisms. There is no logical reason to exclude these possibilities in advance. Put otherwise, without a bona fide theory of cognition, there is no principled reason to suppose that the successes of cognitive psychology depend on mechanisms, and mechanisms alone. To think otherwise would be in the thrall of a fallacious induction. To illustrate: Suppose that all results established in cognitive psychology to date have individuated cognitive activity in terms of mechanisms. Nevertheless, it would be a non sequitur to assume that a more expansive, non-mechanistic mode of individuation would not yield significant and nontrivial results.

Moreover, there are positive reasons to think that there is a need to go beyond the mechanistic and computational, narrowly conceived, and press for a radically enactive or embodied cognitive science. Chemero (2009) provides a clear and compelling case of the need to do so in his argument against neural reductionism that is “based on the details of experimental practice” (p. 170). He cites the unhappy situation in the study of exploratory behaviour in which cognitive psychologists systematically under-described features of object exploration tasks such that their findings, while not useless, were – to use Chemero’s words – potentially confounded. The root problem is that – based on a survey of 116 papers – the majority of work in this area inherited protocols that did not detail the nature of the objects used in their experimental designs to an appropriate degree.⁷ This matters because, as it turns out, the exploratory activity of the species under investigation is greatly influenced by the properties of the objects they deal with and the possibilities for action that those objects afford

them. By ignoring to report these facts the bulk of work in this area failed to “give enough information about the objects used, and so failed to meet one of the primary goals of scientific research: these studies are *not replicable*” (Chemero 2009, p. 173, emphasis original). This is alarming given that molecular neuroscientists, behavioural geneticists and psychopharmacologists have relied on these ‘potentially confounded studies’. This is a clear, yet probably not isolated, case in which the need to go beyond the boundaries of traditional, orthodox cognitive psychology is evident.

It might be thought that proponents of representational accounts of internalism could respond to this problem, at least in principle. For if representations with various contents – contents that represent or stand in for the various affordance-like properties of different objects – feature in cognitive mechanisms then it might be possible to explain how the objects of environmental properties matter without going wide. But, importantly for our discussion here, if one rejects representationalism – as the version of extended functionalism currently under consideration does – this line of reply is closed. Barring representationalism, we have no choice but to accept that specific properties of objects and how organisms relate and respond to them will need to feature in, at least some of, our best scientific explanations of cognitive activity.

Importantly, extended functionalism plays no part in the above argument for blocking and undermining internalist pure functionalism. The last stage of our argument has two main steps. First, it notes that once one goes non-representational there is no clear scientific rationale for, and no clear theoretical means of, thinking of cognitive activity as something smaller or shorter than world-involving extensive relational activity. Second, it provides positive scientific reasons for going wide. Tellingly, the parade case we highlighted from Chemero (2009) works precisely because it rests on a substantial theory from ecological psychology – one that makes appeal to the notion of affordances. It is not a mere negative inversion of internalist functionalism, such that it assumes extended pure functionalism follows if internalist pure functionalism fails. The positive, substantial motivation for going wide in an extensive and not just an extended way does not rest on functionalist considerations. This should hardly come as a surprise for as Putnam explained long ago, when first introducing empirical functionalism, that it was only ever a framework for theorizing and advancing empirical hypotheses; in itself empirical functionalism is “the putting-forward, not of detailed scientifically ‘finished’ hypotheses, but of schemata for hypotheses” (Putnam 1967/1992, p. 54). Yet – and here’s the rub – if extended functionalist does not, by itself, motivate or play any part in convincing us to go wide – if it is only a hollow theoretical frame – we have both the option of dropping it and no obvious reason to retain it. In this light clinging to functionalism in the debates about the extended mind looks like nothing more than intellectual inertia.

5. Extended Functionalism via Commonsense Functionalism

So far so good. But, assuming the above arguments go through, the obvious line of move for the extended functionalist to make is to steer clear of empirical pure functionalism. Not only is this an obvious move – the extended functionalist team à la Clark and Wheeler – have prepared this answer in advance: for their extended functionalism is explicitly advanced under the auspices of commonsense functionalism. In a retrospective on his seminal paper with Chalmers, Clark makes clear his official view on the kind of extended functionalism he endorses “is better viewed as a simple argumentative extension of at least a subset ... [which is non-committal about conscious states] ... of what Braddon-Mitchell and Jackson (2007) describe, and endorse, as ‘commonsense functionalism’ concerning mental states” (Clark 2008, p. 88). Or again, “It is the coarse or common-sense functional role that, on this model (unlike that of empirical functionalism), displays what is essential to the mental state in question” (Clark 2008, p. 88).

Wheeler (2010) adopts the same line: “According to functionalism in the philosophy of mind, ‘what makes something a mental state of a particular type does not depend on its internal constitution, but rather on the way it functions, or the role it plays, in the system in which it is a part’ (Levin 2008). The respective fates of these two positions [extended cognition and functionalism] may not be independent of each other. The claim that ExC [extended cognition] is in some way a form of, dependent on, entailed by, or at least commonly played out in terms of functionalism is now pretty much part of the received view of things [...]” (p. 245).

However, extended functionalism is not, and was not originally thought to be, an immediate consequence of commonsense functionalism. As formulated and advanced under the so-called Canberra Plan that embeds the commonsense functionalism of Lewis (1972, 1995) and Jackson (1998, 2009) and sought to reveal, through conceptual investigations, what the folk ‘find obvious’ about the mental – by attending to and perspicuously representing what lies behind their thought and talk about, *inter alia*, the mental; viz., to reveal the content of our shared implicit folk theory of the mind. So conceived commonsense functionalism only works – it only has legitimacy – if it descriptively captures and states *only* what the folk find obvious about the mental and nothing more. The project depends on accurately characterizing our folk commitments. With this in mind Lewis instructed commonsense functionalists to:

Collect all the platitudes ... regarding the causal relations of mental states, sensory stimuli, and motor responses. ... Add also all the platitudes to the effect that one mental state falls under another ... Perhaps there are platitudes of other forms as well. Include only the platitudes which are common knowledge amongst us: everyone knows them, everyone knows that everyone else knows them, and so on (Lewis 1972, p. 256).

Given this backdrop it is not obvious that unadulterated commonsense functionalism ought to embrace extended functionalism. Consider that opinions are divided about how to respond to Clark and Chalmers' endlessly repeated case of Otto and his notebook. This shows that if we were only to look at the intuitions of the folk, as revealed by how they respond to possible cases, there is no obvious reason to favour extended functionalism over internalist functionalism. In discussing folk reactions to these kinds of thought experiment, Chalmers – in his foreword to Clark's *Supersizing the Mind* – admits that perhaps the opponents of the extended mind thesis “would have commonsense psychology on his or her side. If so, then perhaps this is one point where the ‘commonsense functionalism’ that Clark favors in this book, individuating mental states by the roles that commonsense psychology assigns to them, counts against the extended mind thesis” (Clark 2008, p. xii).

In the light of this, Chalmers recommends the following strategy: “At this point, I think the proponent of the extended mind should not be afraid of a little revisionism. Even if commonsense psychology marks a distinction here, the question still arises of whether this is an important distinction that ought to be marked in this way” (xii). The obvious problem with this suggested move is that to move away from commonsense is to move away from the very project that mandates and motivates commonsense functionalism as described above. In its original formulation the only reason for accepting commonsense functionalism is that it is revelatory of what the folk think about the mental. Without that backing – in pressing for open revisions of what the folk think – extended functionalism would be working without a net. If extended functionalists cannot fall back on an appeal to science and empirical findings (see previous section), it is utterly unclear what could possibly warrant or legitimately constrain the necessary revisions.

Still, it may be thought that extended functionalism's situation is not quite that bad. After all, as Wheeler (2010) readily admits, with respect to traditional variants of commonsense functionalism: “we don't quite have a case of plug-and-play philosophy here. Functionalism – or rather, how we formulate it – *needs to be tweaked a little before current needs are met*. To see why, recall that, according to the traditional formulation of the position as given earlier, a mental state is constituted by the causal relations that it bears to sensory inputs, behavioral outputs, and other mental states. But depending on how one hears terms like “sensory inputs” and “behavioral outputs,” this statement of the view *may harbor a bias toward the inner that isn't, at root, a feature of its defining commitments*. Fundamentally, the functionalist holds that what makes a systemic state a mental state is the set of causal relations that it bears to systemic inputs, systemic outputs, and other systemic states ... Once we have this more general characterization of the functionalist line, we can allow the borders of the cognitive system to fall somewhere other than the sensorimotor interface of the organismic body. And that opens the door to a cognitive system whose boundaries are located partly outside the skin.” (p. 249, emphases added)

Yet this analysis too falls short of what's needed. What extended functionalists need is more than the observation that commonsense tolerates the possibility that cognition might extend or the revelation that commonsense lacks the resources to decide the question either way. To motivate extended functionalism via commonsense functionalism it would have to be established that careful attention to what the folk really think about the mental gives us some substantive, positive reason to believe in extended functionalism.

Where might we look for this? Arguably by examining more closely what the folk do when competently deploying their psychological concepts. However, the best way to do this is to challenge functionalism not to continue to work within its template. For functionalism it itself, arguably, a presumptuous imposition on commonsense (Ratcliffe 2007, Hutto 2011). When we look (rather than 'think' – that is, rather than presupposing and imposing) we find that there are no clear lines in folk talk about the mind that demarcate the boundaries of cognitive phenomena neatly. Folk talk is not designed to tell us where perceptual activity, for example, begins and ends in any precise sense. Arguably what Wittgensteinian and phenomenological investigations into our ordinary thought and talk about the mental reveal is that while the folk do conceive of cognition and perception as kinds of public activity, their spatial and temporal boundaries are messy, rough-edged and extensive (Hutto 2013, Forthcoming). Does this mean the extent of cognition has no limits? Of course not. And those limits can be revealed by empirical experiment. Even if one were able to see for miles one might not be able to see for leagues. There will be empirically discoverable species-wide and individual differences with respect to the limits of cognition – limits that can and should (and are being) actively empirically explored. It is precisely at this interface that philosophy and science can and should fruitfully inform one another.

6. Conclusion

We hope by now to have established that the notion of extensive minds is not hopelessly vague, ill-conceived or, worse, inconceivable. More than that, we have argued that it already figures in explanations required by the sciences of the mind and our ordinary ways of understanding the mental.

As a final word, it should be clear from the discussion above, that endorsing extensive enactivism does not entail leaving neuroscience behind or out of the story. Even if the great bulk of cognitive processes are extensive and world-involving there is still every reason to understand empirically the special contributions that brains make to enabling such cognition. Our motto is: Brains – Don't leave home without them! Still, it only helps, not hinders, neuroscience to make sure the brain is put in its place – for only then will its contribution to cognition be appropriately understood and not unhelpfully exaggerated.

References

- Adams, F and Aizawa, K. (2010). Defending the Bounds of Cognition. In *The Extended Mind*. Menary, R. (ed). Cambridge, MA: MIT Press. 67-80.
- Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, MA, MIT Press.
- Clark, A. (2008). *Supersizing the Mind*. Oxford: Oxford University Press.
- Hutto, D.D. (2011). Presumptuous Naturalism: A Cautionary Tale. *American Philosophical Quarterly*. 48 (2) pp.129 – 145.
- Hutto, D.D., and Myin E. (2013). *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: MIT Press.
- Hutto, D.D. (2013). Enactivism: From a Wittgensteinian Point of View. *American Philosophical Quarterly*. 50: 3. 281-302.
- Hutto, D.D. (Forthcoming). Contentless Perceiving: The Very Idea in *Wittgenstein and Perception*. OSullivan, M and Campbell, M. (eds). London: Routledge.
- Hurley, S. (2010). The Varieties of Externalism. In *The Extended Mind*. Menary, R. (ed). Cambridge, MA: MIT Press. 101-154.
- Hutchins, E. (2014). The Cultural Ecosystem of Human Cognition. *Philosophical Psychology*. 27:1. 34-49.
- Jackson F. 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Jackson F, Mason K, Stich S. 2009. *Folk Psychology and Tacit Theories: A Correspondence between Frank Jackson and Steve Stich and Kelby Mason*. In *Conceptual Analysis and Philosophical Naturalism*, ed. D Braddon-Mitchell, R Nola, pp. 45-97. Cambridge, MA: MIT Press.
- Khalidi, M.A. Innate Cognitive Capacities. *Mind and Language*, 22:1 February 2007, pp. 92–115.
- Lewis D. K. 1972. Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy* 50:249-58.
- Lewis D. K. 1995. Reduction of Mind. In *A Companion to the Philosophy of Mind*, ed. S Guttenplan, pp. 412-31. Oxford: Blackwell.
- O'Brien, G and Opie, J. (2009). The Role of Representation in Computation. *Cognitive Processing*. 10: 53-62
- Piccinini, G. (2008). Computation without Representation. *Philosophical Studies*. 137 (2): 205-241
- Putnam, H. (1967/1992). The Nature of Mental States. In Beakley, B and Ludlow, P. (eds). *The Philosophy of Mind*. Cambridge, MA: MIT Press.
- Putnam, H. (1988). *Representation and Reality*. Cambridge, MA: MIT Press.
- Ratcliffe M. 2007. *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. Basingstoke: Palgrave Macmillan.
- Rowlands, M. (2010). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge, MA: MIT Press.
- Rowling, J.K. (1997). *Harry Potter and the Philosopher's Stone*. London: Bloomsbury Publishing.
- Rupert, R. (2009). *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.
- Searle, J. (1990). Is the Brain a Digital Computer? *Proceedings and Addresses of the American Philosophical Association*. 64: 3. 21-37.
- Sutton, J. (2014). Remembering as Public Practice: Wittgenstein, Memory, and Distributed Cognitive Ecologies. *Mind, Language, and Action: proceedings of the 36th Wittgenstein symposium*, D. Moyal-Sharrock, V.A. Munz, and A. Coliva (eds.). Ontos-Verlag
- Wheeler, M. (2005). *Reconstructing the Cognitive World*. Cambridge, MA: MIT Press.
- Wheeler, M. (2014). The Revolution will not be Optimised: Enactivism, Embodiment and Relationality. *Proceedings of the 50th Anniversary Convention of the AISB*. <http://doc.gold.ac.uk/aisb50/AISB50-S25/AISB50-S25-Wheeler-extabs.pdf>
- Wilson, R. A. (2004). *Boundaries of the Mind: The Individual in the Fragile Science*. Cambridge: Cambridge University Press.

Notes

-
- ¹ To prevent misunderstanding, it is important to clarify what ‘basic’ cognition means on a REC account. It is common in the scientific literature to treat ‘basic’ cognition as if it designates only very low-grade forms of cognition. Yet the intuitive notion of basic cognition at work in the general scientific literature, at best, only accidentally corresponds to the notion of basic cognition used in the REC framework. REC promotes the possibility that there are kinds of mind that lack content, namely that non-representational minds exist. It assumes that such minds are, phylogenetically and ontogenetically, basic. They are the most fundamental kinds of minds. REC also allows that contentful minds exist, but that these are not ‘basic’ in the above sense. Such minds – or better content-involving modes of cognition – are necessarily scaffolded modes of cognition. From a REC perspective contentful thought is not a feature of all cognition, rather it is a special achievement.
- ² Thus Wheeler (2014) claims that “one leading alternative to enactivism in 4E space, namely extended functionalism, is in the position to robustly reject [internalism about cognitive architecture] but, unlike enactivism, has no theoretical interest in rejecting [content-bearing representations]” (p. 1). Indeed he goes further and says, “Against Hutto and Myin, I shall briefly present an analysis which indicates that to the extent that extended functionalism can be adequately defended, it can be defended in a representationalist register” (p. 1).
- ³ Wheeler insists, for example, that explanatory naturalists must satisfy what he calls the Muggle constraint such that: “One’s explanation of some phenomenon meets the Muggle constraint just when it appeals only to entities, states and processes that are wholly nonmagical in character. In other words, no spooky stuff” (Wheeler 2005, p. 5).
- ⁴ As Piccinini notes, “Many ... [will assume that] computational states are individuated by their formal or syntactic properties ... that computational states are individuated syntactically rather than semantically. But this is far from the case” (Piccinini 2008, p. 208). Putting the point more softly, P observes “it has seemed difficult to give an account of syntactic properties without appealing to semantic properties (cf. Crane, 1990; Jacquette, 1991; Bontly, 1998). Nevertheless, he then adds, “I don’t know how to tell whether a property is syntactic” (Piccinini 2008, p. 209).
- ⁵ There is evidence of the inherent circularity in this sort of approach in Rupert’s official formulation: “If there is any theoretically interesting divide between what is distinctively

cognitive and what merely causally contributes to intelligent behaviour, it is to be found in the persisting, integrated nature of *cognitive* architectures” (Rupert 2010, p. 344, emphasis added).

- ⁶ Evidence that such empirical assumptions are being made is easy to find: “We think orthodox cognitive psychology places its bets on the individuation of cognitive processes in terms of mechanisms.” (Adams & Aizawa 2008, p. 125). “We think that the available empirical evidence provides good reason to think that the chances of finding *interesting cognitive regularities* covering brains and tools is low. Bear in mind that we side with that is by all accounts scientific orthodoxy.” (Adams & Aizawa 2010, p. 74, emphasis added).
- ⁷ Chemero (2009) reports that, “the main results of the literature survey were as follows. Of these 116 articles, 52 (approximately 44 percent) gave little or not information concerning the specific objects that were given for exploration, and 64 (approximately 56 percent) provide detailed descriptions of the objects. Of the 64 articles that included descriptions of the objects used in the experiments, 32 (approximately 28 percent of the total) used set of objects with non-equivalent affordances for the species of animal under study” (Chemero 2009, p. 173).