What is left of the active externalism debate?

Victor Loughlin & Karim Zahidi[1]
Centre for Philosophical Psychology
Department of Philosophy
University of Antwerp

Penultimate draft. Accepted for publication by *The European Journal of Philosophy*.

1. Introduction

Since the publication of Clark and Chalmers (1998), "active externalism" has been a hotly debated issue. [2] Defenders of active externalism hold 1) cognitive processes can extend beyond the brain and body and 2) cognitive states can be partly constituted by non-biological, environmental objects and processes. The basic claim is that it is through our active engagement with bodily external elements that certain cognitive processes and states are made possible. As a consequence, cognitive processes and states can be partly *constituted* by features of the world around us. For example, consider the amnesiac, Otto. According to Clark and Chalmers, some of Otto's beliefs include, in an importantly non-trivial sense, information contained in his notebook. Following the literature, we take this as a paradigm case of active externalism.

Detractors of active externalism however challenge both (1) and (2). That is, while they acknowledge that external elements may causally contribute to some cognitive processing, they nonetheless deny that these elements play a *constitutive* role in cognition. So, for example, external props, like Otto's notebook, may help in the performing of certain cognitive tasks, but such props are mere aids for the sui generis cognitive processing that occurs within the head.

[1] Authors are listed in alphabetical order.
[2] Active externalism has traded under many names (environmentalism, vehicle externalism, the hypothesis of extended mind (HEM), the hypothesis of extended cognition (HEC), to pick just a few). For the sake of clarity, and to emphasize that the focus of our argument is directed at the debate as a whole, and not at particular issues within the debate, we shall stick with Clark and Chalmers' original coinage "active externalism".

This critique is further motivated by the worry that calling non-biological objects and processes "cognitive" or "mental" threatens to rob these labels of their metaphysical significance. For if any object or process that causally contributes to the performance of a task is thereby deemed cognitive or mental, then cognition or mentality seemingly spreads arbitrarily into the world. This looks metaphysically suspect, it is claimed, since there then is no principled distinction between those things that obviously have minds e.g. people, animals etc., from those things that obviously don't e.g. rocks, tables etc.

Moreover, active externalism looks scientifically suspect. Consider once again Otto. Otto stores information in his notebook. In order to retrieve this information, he consults his notebook. However, his non-amnesiac counterpart, Inga, retrieves information from her biological memory. Thus, laws governing Inga's biological process will not be applicable to Otto's brain-body-world extended process and vice versa. But if we understand both processes as instances of remembering, then it is unclear how there could be laws under which both processes could be subsumed. However, if the purported goal of cognitive science is the uncovering of such laws, then active externalism looks to undermine this goal.

These familiar criticisms are the standard fare in the literature. Yet rather than zero in on these matters, we propose to do something different, namely broaden the discussion to focus on two questions that seemingly drive or motivate much of this debate.

First, there is the question: in virtue of what is a state or activity "cognitive" or "mental"? Second, there is the question: what, if anything, do the causal mechanisms that underpin or realize our cognitive or mental capabilities have in common?

It is our view that these two questions are separate and should be kept distinct. For while investigating the mechanistic underpinnings of some cognitive or mental capability is no doubt a worthy endeavour, we claim that it will not reveal why such a capability should be characterised as cognitive or mental. This is because the cognitive or mental status of a capability will not be found inscribed within underlying causal mechanisms. In order to identify what constitutes the mental, one has to instead 'go wide', that is, factor in wider social and cultural practices and activities. And while mechanisms do have identifiable spatial and temporal locations, and thus one can ask

whether those mechanisms do or do not extend beyond the body, the same is not going to be true of mentality itself, since mentality is an unbounded phenomenon and so has no location (except in the trivial sense that everything occurs somewhere).

Where does this leave the active externalism debate? If our 'wide view' of mentality is correct, then minds neither extend nor spread arbitrarily into the world. Not because minds are internal but rather because mentality is instead a widely constituted, unbounded and so non-localisable phenomenon. Yet as we will show in this paper, there has been a tendency within the active externalism debate to assume that one can resolve what is a cognitive or mental state or activity by identifying a property or properties common to those underlying causal mechanisms that realize that state or activity. That is, there has been a tendency to try to answer the first question in terms of the second. Contrarily, if mentality is wide, as we claim, then no appeal to underlying causal mechanisms will resolve what is a cognitive or mental state or activity.

However, if the debate were instead to be viewed as committed to *only* answering the second of the two questions, that is, to only focusing on the question whether some extended causal mechanisms on the one hand and those intracranial mechanisms involved in the realization of cognitive or mental capabilities on the other have something in common, then this might be compatible with our wide view. Our wide view then supports re-configuring the active externalism debate such that it only focuses on the second question. Yet we claim that once this re-configuration is fully worked out, then it in fact reveals that the debate as understood by prominent proponents and detractors of active externalism is all but over.

We will proceed as follows. In section 2, we argue that there is - or has been - a tendency within the active externalism debate to assume that the first question can be answered by appeal to the second. That is, some within the debate take the view that the property in virtue of which a state or activity is cognitive or mental is a property or set of properties common to the underlying mechanisms that realize that state or activity. In section 3, we offer our 'wide view', namely the view that the properties in virtue of which an activity is mental are factors of the personal (life-history) and impersonal (socio-cultural) background. Our wide view entails that the two metaphysical questions

previously considered are in fact separate and should be kept distinct. In section 4, we explore what implications this has for the active externalism debate.

2. Active externalism - two questions

Defenders of active externalism hold 1) cognitive processes can extend beyond the brain and body and 2) cognitive states can be partly constituted by non-biological, environmental objects and processes. The fact that *constitution* is the metaphysical term par excellence, combined with the fact that it is used widely throughout this debate, supports the contention that this debate is one primarily focused on the metaphysics of mind. Yet care needs to be taken in determining exactly what is the metaphysical question at work here. For there are (at least) two different questions to which active externalism might be an answer.

For any cognitive or mental state or activity, we can ask:

> **Q1.** In virtue of *what* is a state or activity a cognitive or mental state or activity?

Note that Q1 cannot be answered by simply isolating a property shared by all and only those states or activities that one deems "cognitive" or "mental". This is because in identifying what makes a state or activity cognitive or mental, the aim is to explain *why* that particular state or activity is thereby cognitive or mental.

Yet we can also ask:

> **Q2.** What, if anything, do the underlying causal mechanisms that underpin or realise a given cognitive or mental state or activity have in common?

Note that part of answering Q2 consists in identifying the causal mechanisms that realize a given cognitive or mental activity. This might suggest that answering Q2 is simply an empirical matter. But things are not that simple. Consider the way one might try to understand a cognitive capacity. A frequently used strategy in psychology is to give a high level functional description of the cognitive capacity in question and then

decompose that capacity into sub-capacities, which can then be further decomposed until one obtains a set of sub-capacities that are simple enough to be realized by a set of localizable mechanisms. One way to reformulate Q2 then is in terms of cognitive systems. A cognitive system is a more or less permanent assemblage of mechanisms, which are necessary to perform a cognitive task. If we plug this definition into Q2 , then Q2 becomes: what, if anything, do all cognitive systems have in common? [3]

This is bound up with explanatory practices in cognitive science. For there is the question as to what is the correct high level description of the cognitive activity and what is the correct division of sub-capacities. For example, in order to be able to exercise my ability to remember, various other bodily capacities need to be realized. My blood circulatory system has to be up and running, for instance. But few would want to consider such a system a bona fide part of the human cognitive system, even if it is part of the conditions necessary for cognition to take place. Hence, determining what is common to all cognitive systems requires first distinguishing those mechanisms that are genuine parts of the system from those parts that are not. This is an explanatory issue. For as the example of the blood circulatory system indicates, what we decide to include as genuine parts of such a system will depend on the phenomena we are trying to explain. All of which reveals that Q2, like Q1, is a separate and distinct metaphysical question.

The difference between Q1 and Q2 can perhaps be made clearer by way of a non-cognitive example. Consider the coin in your pocket. We can ask, first, in virtue of what is this coin "money" i.e. what kind of things ensure that this coin has monetary value and is legal tender? But we can also ask: given the fact that the coin is "money", how is that state or activity realized? As we will later argue, to answer the first question one needs to appeal to a number of social, cultural and institutional practices. Yet the second question can be answered by referring to the fact that the coin has the right dimensions, is made out of a certain alloy, etc. This example, which will be discussed further in section 3 of this paper, reveals that answering Q1 or Q2 requires appealing to very different entities and processes.

[3] As we will show (section 4 of this paper), this question is a popular one within this debate.

Yet while these two questions are noticeably distinct, there has been a tendency within the debate over active externalism to assume that answers to both questions can be obtained by solely focussing on the material realizers of cognitive activities. This assumption is far from obvious, however.

In its defence, it might be argued that if one endorses functionalism, then one might think that Q1 is a question about the functional role of the mechanisms while Q2 is a question about the material and causal properties of these mechanisms. To the question "In virtue of what is process Y cognitive?", a functionalist may answer, "Because Y has the same functional profile as process X, and X is cognitive". If we cash out functional profile in terms of the causal properties of the underlying mechanisms, then it would seem that Q1 (for process Y) can indeed be answered in terms of the common feature of the underlying mechanisms that realize X and Y (the common feature being that they are both realisers of the same functional profile).

However, the answer to Q1 for process Y refers, not merely to the properties of the underlying mechanisms, but also to the fact that X is cognitive. Hence, giving a complete answer to Q1 for the process Y forces us to answer Q1 for process X ("In virtue of what is process X cognitive?"). Thus, on pain of infinite regress, if we want a full answer to Q1 for process Y, we will need to answer question Q1 for some process Z without invoking functional profile. That is, invoking functionalism is in itself insufficient to provide an answer to Q1.[4]

Prominent voices in the debate over active externalism have proposed to bridge the gap left by functionalism in answering Q1 by appealing to some (non-functional) feature common to underlying causal processes. That is, they have assumed that the properties

[4] Clark and Chalmers' Parity Principle or PP is a case in point. PP is the hypothetical: if the role played a bodily external process could also be played by a bodily internal process, and we were to recognize such a bodily internal process as a bona fide cognitive or mental process, then we should also accord a similar status to the bodily external process, since, according to our hypothetical, both play equivalent roles (Clark and Chalmers, 1998). PP is now recognized as a functionalist claim (Menary, 2010). Clark and Chalmers' use of PP supports our assessment of the inadequacy of functionalism to answer Q1: application of PP requires that we assume that bodily internal processes are bona fide cognitive or mental processes. But Clark and Chalmers give no indication as to why intracranial processes are in fact cognitive. As such, they do not answer Q1.

or features in virtue of which an activity is cognitive (Q1) are some common (non-functional) feature of the underlying causal mechanisms (Q2).

For example, both defenders and detractors have deployed a mark of the cognitive strategy, namely they have attempted to identify "some necessary conditions on a state or process being cognitive" (Adams & Aizawa 2001, p. 48). According to those who deploy such a strategy, one needs such a mark of the cognitive to resolve the debate over active externalism. Yet there are two possible readings of a mark of the cognitive.

On the first reading, a mark of the cognitive simply seeks to identify what all cognitive processes have in common. It does not ask that the criteria provided specify those properties in virtue of which the processes are cognitive. In other words, this reading begins from the assumption that we already know which processes are cognitive and then proceeds to look for what they have in common. A mark of the cognitive thus understood is a set of criteria that determine the *extension* of the concept "cognitive". In other words, it is a strategy to answer to Q2.

On the second reading however, a mark of the cognitive aims to lay bare the essential core of cognition. That is, it aims to specify those properties *in virtue of which* a process is cognitive. Understood in this way, a mark of the cognitive aims to give the *intension* of the concept "cognitive". On this second reading, the demand for a mark of the cognitive can be understood as equivalent to answering question Q1.

We will argue that those who endorse the mark of the cognitive strategy aim for an intensional mark of the cognitive and so attempt to answer Q1. Yet upon examination of their respective marks, it turns out that they make use of extensional features of cognitive processes, namely they appeal to those properties that the causal processes that underpin a given cognitive activity have in common, that is, they answer Q2. In other words, it is assumed that the answer to Q2 is also an answer to Q1, since it is assumed that the property in virtue of which a state or activity is cognitive or mental (Q1) is a common property of the underlying mechanisms that realize that state or activity (Q2), that is, it is to answer Q1 via an appeal to Q2.

For example, Adams and Aizawa, who are well known detractors of active externalism, are not always clear about what question their mark of the cognitive is supposed to answer. They initially introduced their mark as follows:

> "A first *essential* condition on the cognitive is that cognitive states must involve intrinsic, non-derived content." (Adams & Aizawa 2001, p. 48, our italics)

The use of a term like "essential" suggests that they are aiming for an intensional mark of the cognitive, that is, they aim to specify in virtue of what is a state cognitive. In later work, this motivation is made clear:

> "It is easy to give a list of cognitive processes. They are things like learning, memory, concept formation, reasoning, maybe emotion, and so on. It is not easy to say, of these things that are called cognitive, *what makes them so."* (Adams & Garrison 2013, p. 340, our italics)

It is this difficulty that then sparks the need for a mark of the cognitive. A mark of the cognitive is thus an attempt to answer Q1.

However, the essential condition of a mark of a cognitive, according to Adams and Aizawa, is that it involve processes that possess intrinsic or non-derived content:

> "cognition is constituted by certain sorts of causal processes that involve nonderived contents." (Adams and Aizawa, 2010, p. 68)

Thus, on the one hand, Adams and Aizawa seemingly acknowledge that there are two questions at work here, namely one to do with what makes a process cognitive, the other to do with what feature the causal mechanisms that underpin cognitive capabilities have in common, viz. non-derived content. [5] On the other hand, they also assume that

[5] Non-derived content is informational content whose existence *qua content* is not dependent on the existence of other informational content. The standard examples are the content involved in an agent's own thoughts, perceptions, experiences etc. Derived content, on the other hand, is informational content whose existence *qua content* does depend on other content. Here the standard example is public language.

what determines the cognitive status of a process is a characteristic property common to the mechanistic base of that process, viz. non-derived content.

In recent work (Adams & Aizawa, 2010b), the idea that cognitive status is determined by the intrinsic properties of the underlying mechanisms is made explicit. Adams and Aizawa argue that even if one doubts the coherence of the notion of non-derived content, this in itself would not mean a victory for active externalism, since rejecting the notion of non-derived content,

> "leaves untouched the idea that part of what distinguishes cognitive processes from non-cognitive processes is *the way in which* cognitive processes transform or manipulate mental representations. Cognitive processing is not just any old type of symbol manipulation or information processing." (Adams & Aizawa 2010b, p. 581, our italics)

Again, this makes clear that Adams and Aizawa aim to offer an intensional mark of the cognitive, since they aim to resolve what a cognitive process is. As such, their mark is an attempt to answer Q1. But as we have seen, their mark in turn involves appeal to extensional features, that is, to that property which, they allege, all cognitive processes have in common, namely non-derived content. Yet this is to assume that the properties or features in virtue of which an activity is cognitive (Q1) is some common (non-functional) feature of the underlying causal mechanisms (Q2). That is, it is to assume that Q1 can be answered via Q2.

The idea that a mark of the cognitive can be located in the property or properties common to those processes that underpin or realize cognitive or mental states can also be found in Rowlands (2010). Contrary to Adams and Aizawa, Rowlands is a prominent defender of active externalism (or what he calls "Amalgamated Mind"). However, like Adams and Aizawa, Rowlands is also convinced that one needs a mark of the cognitive to resolve the question as to whether (some version of) active externalism is correct. He writes:

Public language has content, but that content is dependent on conventions, such as socio-cultural practices, whose existence in turn depends on the thought-contents of individuals.

> "it is not possible to establish the cognitive status of the latter [bodily external processes] simply by *analogical extension* from the former [bodily internal processes]. Therefore, if we are to defend the cognitive status of the extended processes, we need some other way of defending the claim that these external processes are cognitive ones. One way of doing this — and it is not clear that there is another way — would be to provide an adequate and properly motivated criterion or mark of the cognitive: a criterion that would allow the thesis of the extended mind to justify the claim that the external processes involved in cognition are indeed cognitive processes." (Rowlands 2010, p. 90, our italics)

The question can then be asked: does Rowlands understand a mark of the cognitive as a demand for an extensional set of criteria (the first reading) or as an attempt to lay bare the essential core of cognition (the second reading)?

In the above passage, Rowlands insists that we cannot reason by analogy to confer cognitive status to extended processes. This arguably militates against the (first) extensional reading of the mark of the cognitive. For suppose we identify those features that all cognitive processes we have previously encountered have in common. Reasoning analogically, we could then say that extended processes i.e. ones involving an agent and an environmental object or process, are cognitive if we find that such processes share those features. Yet Rowlands (correctly in our opinion) makes it clear that this way of proceeding is invalid. With his mark of the cognitive, he instead aims to identify the way in which a process is cognitive. As such, he is committed to the second reading, that is, he aims to offer an intensional mark of the cognitive.

This is further corroborated by the fact that Rowlands claims that active externalism is an "ontic" claim, as opposed to merely epistemological claim, that is, it is a claim about "what (some) mental processes *are*" (ibid, p. 59, our italics). Hence, like Adams and Aizawa, Rowlands' mark of the cognitive is meant to answer question Q1.

Rowlands' answer to Q1 is given by the following four sufficient conditions:

> "A process P is a cognitive process if:
> 1. P involves information processing — the manipulation and transformation of information-bearing structures.
> 2. This information processing has the proper function of making information available either to the subject, or to subsequent processing operations within the subject; information that was, prior to this processing, unavailable.
> 3. This information is made available by way of the production, in the subject of P, of a representational state.
> 4. P is a process that belongs to the subject of that representational state." (Rowlands, 2010, p110-111)

These conditions reveal that, even though there are differences between Rowlands' mark of the cognitive and that provided by Adams and Aizawa, both share a common core. [6] This is especially clear if we consider conditions (2) and (3) in conjunction with Rowlands insistence that representations mentioned in (3) must have non-derived content: taken together they seem to form a paraphrase of the Adams and Aizawa's mark, which states that a causal process is cognitive if it involves non-derived content. Yet if so, then Rowlands' mark is simply an elaboration of Adams and Aizawa and the criticism that we levelled at the latter's mark also applies to the former's. In particular, on the one hand, Rowlands' aims to offer an intensional mark of the cognitive, that is, to identify what makes a given process cognitive and so answer Q1. But, on the other hand, he does so in terms of extensional features, namely in terms of what all such processes have in common (namely his four conditions). Thus, like Adams and Aizawa, Rowlands' mark assumes, without argument, that the property that makes a state or activity cognitive (Q1) is a property (or a set of properties) of the underlying causal

[6] One such difference is that Adams and Aizawa offer their mark as a necessary and sufficient criterion, whereas Rowlands offers his mark as a set of merely sufficient conditions. Rowlands thus allows for the fact that there might be cognitive processes which do not meet the four criteria making up his mark of the cognitive. However, this difference is compatible with our point, namely that Rowlands, like Adams and Aizawa, assumes that Q1 can be answered via Q2. For instance, when defending his mark, Rowlands appeals to the fact that all cognitive processes hitherto discovered satisfy his four criteria. Rowlands thus views his mark as extensionally adequate but, in contradistinction to Adams and Aizawa, refuses to reason analogically to conclude that all cognitive processes have to satisfy the four conditions. But then it follows that his mark cannot give the intension of the notion of cognition. Yet as we point out, Rowlands nonetheless views his mark as giving such an intension of cognition. Hence it appears Rowlands wishes to give intensional account of cognition (Q1) via an extensional account of cognition (Q2). This is to assume that Q1 can be answered via Q2.

mechanisms (Q2). In other words, like Adams and Aizawa, Rowlands attempts to answer Q1 via an appeal to Q2.

We have now shown that both detractors of active externalism (Adams and Aizawa) and defenders of active externalism (Rowlands) assume that Q1 can be answered by appeal to Q2. Why is this assumption problematic? We think it problematic because, prima facie, Q1 and Q2 are in fact separate and distinct matters. Hence treating them as merely different versions of the same underlying problematic is unwarranted, unless one provides a convincing argument as to why such an assumption is harmless or insubstantial.

Yet we think there are good grounds to claim that Q1 and Q2 are independent, that is, answering one question does not constrain how to answer the other. In the following section 3, we will defend the position that mentality is constituted by 'wide' social and cultural practices and activities. This 'wide view' reveals that identifying what constitutes the mind (e.g. wide practices and activities) is a separate and distinct matter from resolving issues concerning the properties common to the mechanistic base for a given cognitive activity. In which case, the issues raised by Q1 cannot be resolved by consideration of the issues raised by Q2. We thus think there is reason to challenge the assumption that Q1 can be answered via Q2. [7]

3. Going Wide

3.1 The Muggle Constraint

[7] Some may object that all we have offered so far is an interpretation of certain writings and such interpretations need not correspond with the author's true intentions. Others may object that we have only shown that some participants to the active externalism debate endorse the offending assumption, not that all do. We concede both points. Our interpretation could be counter-posed with another interpretation, which perhaps may not display the same problematic assumption concerning the relationship between two different projects, viz. answering Q1 and answering Q2. Further, other participants to this debate may not endorse the assumption we have identified. However, even if such a rival interpretation of Adams and Aizawa or Rowlands could be offered, and even if it is not clear that other participants to this debate do assume that Q1 can be answered via Q2, these facts would still support our main point, namely that the distinction between the two questions has not been sufficiently attended to within the debate.

We have argued that active externalism can be interpreted as answering two distinct metaphysical questions. We have claimed that there has been a tendency within the debate to assume a tight relationship between these two questions. This is manifested in the idea, endorsed by prominent voices on both sides in the debate, that the property in virtue of which a state or activity is cognitive or mental is a property common to those underlying mechanisms which realize that state or activity.

We however deny that the mental status of an activity is in fact inscribed within the mechanisms that underpin or realize it. Framed in terms of the two questions, it is our view that even if there are properties common to the causal mechanisms that underpin some mental activity (Q2), such properties will not reveal what makes such an activity mental (Q1). Q1 cannot be answered via Q2. If one wants to identify what makes an activity mental i.e. one's aim is to answer Q1, then one needs to 'go wide'. This is our 'wide view' of mentality. On this view, what constitutes a given mental activity are those features of the environment – whether physical, biological or socio-cultural – within which such activity is situated. Skipping ahead slightly, our full answer to Q1 will be that an activity is mental when it sustains and explains intelligent behaviour, which it will do via a suite of wide factors, factors that will (1) cover a multiplicity of practices and activities and (2) explain intelligent behaviour in different ways. Mental activities are, in short, widely constituted. As such, the properties of underlying causal mechanisms are not in and of themselves that which makes a given activity mental. From our perspective then, question Q2 is irrelevant for answering question Q1. We neither deny or affirm that there might be something in common between all mechanisms that realize cognitive activity – allthough on the wide view that seems rather unlikely

Before defending our wide view, it is worth first addressing the following issue. It may strike some as unscientific to claim that the properties in virtue of which an activity is mental are not the intrinsic properties of the underlying mechanisms. Does this not run counter to naturalism? And if it does, how could a non-naturalist account be reconciled with good scientific practice? These questions have particular resonance within the debate over active externalism, since active externalism is generally understood to be a naturalist claim. Thus, if challenging the debate about active externalism were to be motivated by non-naturalist considerations, then the defender of active externalism

might object that we have begged the question. For if one already endorses a non-naturalist account of mentality, then one is not likely to think active externalism true.

This is a legitimate concern. Nonetheless, we maintain that our wide view is fully compatible with naturalism. A naturalist account, by our lights, is one that satisfies what Wheeler nicely titles "the Muggle constraint", namely

> "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 allowed." (Wheeler, 2005, p5)

It is our view that going wide about mentality, that is, appealing to the wider socio-cultural environment, does not introduce any "spooky stuff" and as such is fully compatible with a naturalistic framework. Admittedly this is to break ranks with the sort of austere naturalism that is presupposed by many currently working within the philosophy of mind. Under such an austere view, naturalizing the mind means explaining mental phenomena only in terms of mechanisms recognized by the basic natural sciences (physics, chemistry, biology, neuroscience etc.). Alternatively, we endorse a much more "relaxed naturalism" (Hutto & Satne, 2015), one which widens its explanatory basis to include all of the empirical sciences, including anthropology, sociology etc.

In order then to set the scene for our wide view, we will begin by considering a non-mental example, namely money. What makes any token metallic coin money? One might reply that the coin is legal tender because of its material properties. Yet apart from the fact that coins come in a huge variety of forms, are made up of many different types of alloy etc., one would expect that at least some of the properties that constitute the money-character of the coin are also found in other forms of money, such as paper money or virtual money. But paper money has a different material substrate from coins. And virtual money has no material substrate at all. Clearly, something other than material properties must be what matters.

A more promising reply might come from looking at the causal history of the coin. On this view, a coin is legal tender because an institution that is certified to issue it has produced it. This means that the metallic coin has been produced via a certain physical process, which is approved by that institution and wider society. Counter-factually, if the very same coin had been physically produced on the very same machine but not by the required institution, then it would not have the required money status. If so, then it is not simply the causal history of the coin per se that constitutes its money character. Rather it is the causal history embedded within a specific socio-cultural (economic, political, institutional etc.) matrix. Moreover, the money character of the coin will also be constituted by how it is subsequently used during a wide range of socio-economic transactions. As Pierre Vilar remarks with respect to paper money and coins:

> "This is all 'fiduciary' money, because its use as payment is due to the decision of an institution, and continues as long as the public has confidence either in the ability of the issuing authority to cash it (as with a currency 'convertible' to gold), or in the stability of the buying power it represents (if it is not convertible to gold)." (Vilar 1984, p. 20)

These transactions (buying and selling, converting money to gold) are themselves dependent on larger social, cultural and even political norms. If correct, then the constitutive properties of the coin i.e. those properties that make a token metallic coin "money", will include physical processes and wider socio-cultural practices. We term such constitutive properties "wide".[8]

Wide properties (physical, social, cultural, normative etc.) play causal roles in the sense that they determine a whole background into which the behaviour or use of certain items is situated. Continuing with our money example, suppose I go into a shop to purchase an item. I hand money to a seller and in return I receive an item. Why am I the "buyer" and he/she the "seller"? Why does the token/paper/credit card I hand to the seller count as "money"? Why would my simply taking the item and leaving the shop count as "stealing", "theft"?

[8] We call a feature F constitutive for property P iff something that has property P, has that property (partly) in virtue of displaying feature F.

These questions, we claim, can only be answered by appealing to the relevant social and cultural practices. This reveals wide properties to be diachronic, that is, extended in time and space, across multiple actions and interaction. Such diachronic properties are the background within which terms like "shop", "buyer", "seller", "credit card", "stealing" etc. have purchase. This is not to deny that synchronic properties matter. After all, I have to hand the token/paper/credit card to the seller in order to buy the item. Synchronic properties however are not constitutive. That is, such properties are not what make my acquiring the item in the shop one of "buying" and not "stealing". But if so, then the constitutional properties of coins cannot be understood solely in terms of those narrow causal processes in which a specific coin partakes. Importantly, this does not require appealing to any "spooky stuff". As such, money can be viewed as widely constituted without renouncing naturalism.

Detailing this non-cognitive example of money was important for two reasons. First, it helps clarify what we mean when we talk of wide properties. Second, it shows that talking about wide properties is naturalist (in our relaxed sense).

In the following sub-sections, we will discuss two explicitly cognitive examples, namely writing and memory. These examples are particularly pertinent to discussions about active externalism, since claims about writing and memory have been used to both vindicate active externalism (e.g. Clark and Chalmers 1998) and, in the case of memory, refute active externalism (e.g. Adams and Aizawa, 2001; 2010; Rupert, 2004). Following our previous discussion, we will refer to all properties that are not properties of the mechanistic base of a given mental or cognitive process or state as "wide". As should now be clear, these wide properties will include the impersonal background (e.g. the social and cultural background) but they can also involve more personal factors (personal idiosyncrasies, personal past history etc.).

3.2. Writing

Suppose you or I were to observe two people making inscriptions on a piece of paper. The first person is clearly writing a letter. The second person is making random inscriptions on paper (perhaps due to some uncontrollable movements of their hand).

What makes the activities of these two people different? For starters, we can see that the first person is writing. That is, her actions accord with the sort of practices and norms that we, the larger linguistic community, call "writing". For example, the marks she has made on the page qualify as words. These words are formed into sentences. If asked, she can read to us what she has written. If we want, we too can read what she has written. We are thus prepared to say that she is "writing a letter". By contrast, the actions of the second person do not have any of these distinctive characteristics. On no possible interpretation, can the second person's actions be regarded as "writing".

This distinction between the activities of these two people is, we claim, made possible because of wider socio-cultural practices and norms. This is not to say, for example, that there will be no important differences in brain activity between these two persons. The second person may indeed be suffering from a physical deficit as the result of an injury to their brain. Nonetheless, these differences in brain activity are not what explain why only one of them is engaged in writing. Explaining this difference requires recognizing that it is practices and norms, and not, say, brain activity, which ensure that the actions of the first person count as writing but not the actions of the second.

We take this as grounds to claim that writing is an example of a cognitive activity that is widely constituted. That is, it is constituted by properties involving social, cultural and linguistic practices and activities. If so, then a question like Q1 (what is writing?) cannot be answered by appeal to question Q2 (what is common to the causal mechanisms ongoing, in our heads or elsewhere, when you or I write?). On a 'wide view', these two questions are revealed to be separate and distinct.

Now, this view might be challenged in two ways. First, it might be objected that wide properties, like practices and norms, are simply 'enabling' conditions. That is, they are the conditions that need to be in place for one to designate certain actions "writing". However, none of these enabling conditions, according to this objection, are in virtue of which a certain activity is a token instance of "writing". What makes a certain activity "writing" are instead the proximate causal processes (in the head or elsewhere).

Second, one might claim that writing is not itself a single act. For example, writing involves forming intentions about what to write, perceiving what you've written,

deciding how to continue with the next word etc. While it could be true that wider practices and norms may bind all these different activities together, the cognitive status of these various activities is in fact inherited from the different sub-processes involved. In which case, it is these sub-processes that constitute writing, not wide factors.

While these objections are worth considering, we don't find either of them convincing. For example, we think enabling conditions *are* constitutive. Consider the following variant to our writing example. Imagine the same two people as before: one is writing a letter; the other is making random inscriptions on a page. But now also imagine that we completely strip away all background, practices and norms etc. Clearly, we would have no basis upon which to designate the actions of the first "writing a letter". For while there would still remain important synchronic causal differences between these two people (in their brains or elsewhere), the explanatory reasons for calling the actions of the first "writing" and refraining from applying such a designation to the actions of the second would be entirely absent. This then supports our contention that the enablers of writing, such as wide properties, are in fact constitutive, since without such properties it is unclear how any activity could be designated as "writing". This handles the first objection.

In response to the second objection, we think it simply begs the question. The second objection views writing in terms of intentions, perceptions, decisions etc. These various acts are then understood as affairs solely internal to the agent. That is, the agent has the intention, perception, or the agent makes a decision, and then the agent writes. If so, then the constitutive properties of writing - what makes the subsequent act "writing" - will all be internal to the agent, since the actual pen-to-paper activities are what merely facilitate the externalization of this internal process. In other words, it is the particular properties of some synchronic causal processes (in the head, say) that make a given activity "writing". Yet this is to assume that questions about the constituents of writing can be answered by identifying those properties common to the causal processes that underpin a given act of writing. That is, it is to assume that the property in virtue of which an activity is "writing" (Q1) is a property common to those underlying causal mechanisms that realise that activity (Q2). However, given that we aim to challenge the assumption that Q1 can be answered via Q2, it can be no objection to our wide view to simply assume that they can.

Of course, demonstrating that this second objection begs the question is not enough to vindicate our wide view. In order to further substantiate our wide view, we turn now to another cognitive example, namely memory.

3.3. Memory

What constitutes memory? [9] As Clark and Chalmers' example of the slightly amnesiac Otto reveals, this question has been central to the debate over active externalism. Yet both detractors and defenders of active externalism agree that remembering is that activity by which a subject, through their own actions, makes information that was somehow stored in some medium conscious. Where they differ is on what kind of activity and what kind of storage needs to be involved to call an activity one of "remembering". For internalists, like Adams and Aizawa, the medium where memories are stored is the brain, and the activity that brings the information to consciousness is essentially internal. For externalists, like Clark and Chalmers (and others), the medium and hence the activity of retrieving, may involve bodily external objects such as notebooks and other props.

This reveals that underpinning both internalist and externalist views of memory is the assumption that "recall" should be understood as synonymous with "information retrieval". For once we equate "recall" with "retrieval", then it follows that there must be a repository of memories and a mechanism by which we retrieve these stored memories. Understanding memory then requires knowing what kind of thing the repository is and how the information gets stored there. Indeed, only if we understand 'memory-as-information-retrieval' does the question as to whether the repository of our memories is in or outside the head have any purchase. Yet we can ask: is this assumption accurate? [10]

---

[9] The following is based on Myin and Zahidi (2014). See also Moyal-Sharrock (2013).

[10] The Merriam-Webster dictionary views "to get information from a computer or disk" as a secondary meaning for "to retrieve". Recall and information retrieval may indeed be synonymous terms when it comes to understanding the operations of a computational mechanism. However, much further work (to say the least) is needed to show that what applies to a computer is also applicable to human memory.

Let's take it as given that memory is indeed the recall of things from the past to the present. Now consider the following example: I recall the fact that yesterday at 3pm I was preparing a lecture. In a loose sense, one can say that my recalling this event involves information e.g. information about what I was doing yesterday. But that does not mean that I stand in a certain relation, let alone a spatial relation e.g. retrieval, access to, to a mental vehicle of information. In fact, my remembering what I did yesterday involves much more than simply having access to information.

For example, a crucial condition is that this information needs to be *entrenched*. This notion of entrenchment has various aspects. One such aspect is *contextualisation*. In order for information to be remembered, the subject has to be able to insert the remembered event within a diachronous series of events, which together form part of his or her past life. Thus, to say that I remember what I did yesterday at 3pm is not only to be able to say that I was preparing a lecture. It is also to be able to connect this event to other events in which I was engaged before and after. For instance, my ability to remember what I was doing yesterday could be based on the fact that preparing a lecture is what I normally do at 3pm on a Monday afternoon. Thus, the remembered event is remembered precisely because it is appropriately contextualised within a temporal framework. [11]

These sorts of considerations lead Myin and Zahidi (2014) to conclude that memory is not

> "an archive of representations to which one gains access. Remembering, on the contrary, becomes a capacity to show behaviour in certain contexts in which the influence of previous interactions with the environment can be discerned. […] Such a nonrepresentational conception of memory allows us to spread the context of remembering in time, in contrast to an "object-oriented" conception (Hutto 2006), as when memory is seen as information

[11] We can also point to other forms of entrenchment, such as the fact that memories are deeply interwoven with emotions. One can, for example, be reproached for not remembering the romantic appointment with a lover as being a sign of disengagement. This reproach would be meaningless if emotions were irrelevant to the memory. There are further indications from psychology that memory is to be seen as an aspect of the broader category of mental time travel.

> carried by vehicles or traces. An embodied approach does not link remembering to the discrete events of creation, storage and retrieval of "a memory", but rather considers remembering to be a consequence of a number of not precisely datable converging processes and tendencies, none of which individually constitutes such a discrete event, but is only a necessary, but not sufficient condition. Already existing habits, exogenously induced or endogenously noticed salience, either explicit or implicit, are not discrete features, but related to attitudes of a person, without precise temporal boundaries, but nevertheless, as was illustrated earlier, necessary conditions for later remembering." (Myin and Zahidi, 2014, p. 404)

There are thus grounds to challenge the 'memory-as-information-retrieval' view. Yet since both internalists and externalists are committed to this 'memory-as-information-retrieval' view, then challenging that view supports the contention that the debate between internalists and externalists over memory is in fact based on a *misunderstanding* of the human cognitive capacity of memory.

This is not to deny however that some instances of remembering may in fact involve the retrieval of information. To borrow Clark and Chalmers' famous example, if Otto consults his notebook to find the address of MoMa, and then remembers the address of MoMa, then clearly his retrieval of that information was instrumental to the exercise of his capacity to remember. What we deny however is that it is in virtue of retrieving information that Otto's activity is one of remembering. Note that this is not to support internalism. First, the internalist denies that Otto is remembering on the grounds that Otto is accessing information in an external medium. But, second, the internalist accepts that if the same information had been stored internally, then Otto would remember. Challenging the 'memory-as-information-retrieval' view entails denying both claims. Instead, if Otto's activity is to be called one of "remembering", then this is because such activity is appropriately entrenched, contextualised, and situated within Otto's wider mental life.

This is still to allow that all kinds of causal processes, both bodily internal and bodily external, may be necessary for the capacity to remember. Studying these processes may

be of the utmost importance to understand what causes memory disorder. But this does not entail that studying these processes can shed light on what constitutes memory. For once the 'memory-as-information-retrieval' view is abandoned, then the capacity of memory can be understood to be always an entrenched, contextualized capacity. These considerations make clear that to answer the question, "what makes it the case that a person is remembering something at time t?", one has to appeal to a wide variety of properties that go beyond those processes which happen to be ongoing at time t, either in the brain of the person, and elsewhere. We call this a wide view of memory.

This is not simply an appeal to mental holism, however. To remember also involves an active relationship with the present environment. Remembering, for example, can involve looking for clues in the present environment that help you recall a certain fact or situation. In doing so, we make use of markers, which might involve the bringing into being items that were not already part of our environment e.g. Otto's inscriptions in his notebook. Or it might consist in explicitly relating certain pre-existing objects with certain actions e.g. in wandering through an unknown city, I might say to myself, "at the junction with the bookshop, turn left". These activities may form a part of my ability to remember at a later time.

If these practices are viewed as a part of remembering, then it becomes clear that remembering is co-constituted by socio-cultural norms. That is, *what* is remembered is subject to socio-cultural norms e.g. one should remember the birthday of a loved one, as is *how* things are remembered. For example, instead of simply making a mental note "turn left at the bookshop", I could paint in big letters on the window of the bookshop the instruction "TURN LEFT". But that would not be considered an exercise of my later ability to remember. It would more likely be viewed as an act of vandalism. Socio-cultural norms are thus arguably constitutive of what does and what does not count as remembering.

Thus, what makes an act one of remembering (Q1) is determined not by the properties common to those causal processes or mechanisms that underlie the act itself but rather by how the act is related to wide (social, cultural, impersonal, personal) factors. In which case, Q1 (what constitutes memory?) and Q2 (what are the properties common

to the causal mechanisms that underpin or realize the human capacity to remember?) are revealed to be separate and distinct questions.

3.4 The wide view of cognition

We have argued that two prime examples of cognitive activities are constituted by wide phenomena. However, someone might object that we still have not answered the crucial question, namely in virtue of what are these various activities "cognitive"? Since we deny that activities are cognitive in virtue of some feature common to the causal mechanisms underlying or realizing these activities, then an alternative characterization needs to be given. And to prove our main point i.e. that cognition is widely constituted, this characterization should not rely on features, common or otherwise, of causal mechanisms.

One can of course define the notion of cognition in any number of ways. However, as stated (see section 3.1), active externalism is generally understood to be a naturalist claim, one based on the current state of play within the cognitive sciences. Thus, if our wide view is to critically engage the active externalism debate, then our wide view needs to support an understanding of cognition that is sensitive to the way the term is used in the cognitive sciences.

A brief survey of the literature shows however that there is no full-fledged unitary notion of cognition at work in cognitive science. For example, the notion of cognition employed in ecological psychology or in the dynamical systems approach to cognition differs in important respects from that used in computational cognitive science. Hatfield (2014) argues convincingly that any notion of cognition that is broad enough to accommodate the variety of notions in the cognitive science literature has to be defined at the *molar level*, i.e. at the level of the behaviour of the whole organism. Hatfield takes the notion of intelligent behaviour to be the key notion in terms of which cognition has to be defined.[12] Behaviour is intelligent whenever it is "adaptive, appropriate, and

[12]This is similar in spirit to some of Clark's pronouncements, e.g.: "What makes a process cognitive, it seems to me, is that it supports intelligent behavior. This is obviously unhelpful, though it is almost certainly just the reply that would be given by, say, the average neuroscientist or cognitive psychologist. […]. To identify cognitive processes as those processes, however many and varied, that support intelligent behavior may be the best we can do" (Clark 2010, p. 92-93).

flexible in relation to environmental and organismic circumstances" (Hatfield 2014, p. 367). Cognitive activities are then defined as those activities that sustain and explain intelligent behaviour.[13]

Clearly activities such as remembering and writing are often invoked in the explanation of intelligent behaviour. But notice, first, the multiplicity of wide factors at work here. As we have shown, such factors will include practices, such as linguistic practices and activities, like using notebooks to remember locations, or contextualizing information, that is, fitting information within a diachronous series of events. Notice also, second, that writing and remembering can explain intelligent behaviour in different ways.

Take the case of writing. We see someone scribbling away on a piece of paper. Is this intelligent behaviour? If we see that the person is indeed actually writing (perhaps they are writing to their boss to explain why they are quitting their job), then that behaviour will usually be classified as intelligent. If someone presses further and asks why this counts as intelligent, we will explain what writing is (appealing to the sort of wide factors discussed in section 3.2, like linguistic practices). Now take the case of remembering. Remembering that I gave a lecture last Monday can explain why my behaviour is intelligent because, say, such remembering reminds me to prepare for next Mondays lecture. If someone presses further and asks why this counts as intelligent, we will explain what remembering is (appealing to the sorts of wide factors discussed in section 3.3, activities such as using notebooks to remember or contextualising information).

Hence, on this account, writing or remembering are cognitive activities, not because of some property common to the underlying mechanisms that realize the activity, but in virtue of the fact that they explain why some behaviour is intelligent behaviour. Furthermore, in explaining how writing or remembering functions and contributes to intelligent behaviour, we appeal to a multiplicity of wide factors, hence writing or remembering, qua cognitive activity, are widely constituted. Thus, our answer to Q1 is

[13] Where we distance ourselves from Hatfield's account of cognition is on his insistence that the activities should "involve mentalistic notions such as representations and mental content". This strikes us as an ad hoc restriction on the notion of cognition. Indeed, as Hatfield himself acknowledges, this fails to accommodate the notion of cognition as employed by non-representational accounts of cognition.

that an activity is mental when it sustains and explains intelligent behaviour, which it does via a suite of wide factors. And if at least some activities are cognitive in virtue of wide factors, then the contrary idea that all activities are cognitive in virtue of a common feature of the underlying causal mechanisms must be false. We call this a wide view of mentality.[14]

4. A future for Active Externalism?

In section 2, we claimed that active externalism could be viewed as an answer to two separate and distinct metaphysical questions (Q1 and Q2). We pointed out however that there has been a tendency within the debate to assume that the debate over active externalism is one that requires resolving Q1 via an appeal to Q2. In section 3, we outlined our wide view of mentality. If our wide view is correct, then Q1 and Q2 are separate and distinct questions, since a wide view e.g. of writing or memory, entails that questions about what constitutes cognition (Q1) cannot be answered by identifying the property (or properties) that are common to the causal mechanisms that underpin or realise a given cognitive state or activity (Q2). What implications does our wide view have for the active externalism debate?

Consider one final time Clark and Chalmers' example of Otto and his notebook. Our wide view entails that if Otto's activity i.e. his use of his notebook, is to be understood as mental, then this is not because of a property of Otto's brain or body, or even a

[14] One might object here that the property common to the underlying causal mechanisms could be that they all "contribute to intelligent behaviour". This would answer Q2. However, if so, then Q1 (in virtue of what is a state or activity cognitive or mental?) can be answered via Q2 (what, if anything, do the underlying causal mechanisms that underpin or realise a given cognitive or mental state or activity have in common?), since the property common to the underlying mechanisms does reveal in virtue of what is a state or activity cognitive or mental, namely the property "contribute to intelligent behaviour". This would then challenge our claim that Q1 cannot be answered via Q2. We don't find this objection convincing, however. The objection claims that the property common to the underlying causal mechanisms is that they "contribute to intelligent behaviour". However, it is unclear how this description could be translated into a property of the underlying causal mechanism, for this description in no way constrains the causal realizers of the intelligent activity. (Different causal realizers contribute in different ways, e.g. instances of remembering and imagination both contribute to intelligent behaviour, but nobody would suggest that they contribute in same way and therefore that the causal realizers of these acts of cognition share some relevant property). As such it seems difficult to construe this as a genuine property of the causal mechanisms. Contrast this with the property "realizing the same functional profile". This property does constrain the make-up of its causal realizers, hence in this case functional profile can be seen as genuine property of its realizers. Since description "contribute to intelligent behaviour" is not a property of the underlying mechanisms, our appeal to "contributing to intelligent behaviour" in describing cognition does not imply that we answer Q1 in terms of Q2.

property of a particular process in which Otto is currently engaged at time t. It is not even because of a property that is precisely localisable in space and time. Rather, Otto's use of his notebook is mental because it sustains and explains Otto's intelligent behaviour. And his use of the notebook sustains and explains his intelligent behaviour in virtue of wide factors, that is, in virtue of the personal (life-history) and impersonal (socio-cultural practices) background against which Otto's activities play out.

Yet both defenders and detractors of active externalism assume that what makes a state or activity cognitive or mental is localizable in space and time. For active externalists, like Rowlands, such processes span brain, body and world at time t, or some relatively narrow time-interval around t. For internalists, like Adams and Aizawa, it is a set of properties of brain processes at time t, or some relatively narrow time-interval around t. Our assessment of the active externalism debate explains why they make this assumption. For once it is accepted that determining what is a cognitive or mental state or activity (Q1) can be resolved by identifying those properties common to the causal mechanisms that underpin or realize a given cognitive or mental capability (Q2), then it follows that cognition or mentality is something locatable, since causal mechanisms obviously have identifiable spatial and temporal locations.

Yet if our wide view of cognition is correct, then mentality is an unbounded phenomenon and so is not localisable in space and time (except in the trivial sense that everything must occur somewhere). [15] Our wide view thus entails that both internalists and externalists are mistaken. In other words, if the debate over active externalism is understood as an attempt to answer Q1 via an appeal to Q2 (as both the Adams and Aizawa and Rowlands' marks of the cognitive suggest it is), then going wide about mentality entails that active externalism as understood by prominent voices on both sides of this debate requires dissolution, not solution. [16]

[15] We are not alone in thinking of mentality as unbounded. Others who endorse similar but nonetheless distinct positions are, for example, Bennett and Hacker (2007), Thompson and Stapleton (2009) and Hutto and Myin (2013).

[16] Active externalism (Clark and Chalmers, 1998) was originally framed in terms of cognitive processes e.g. the parity principle, and cognitive (or mental) states e.g. beliefs. However, if mentality is wide, as we claim, then mentality should instead be framed in terms of behaviours, activities, practices, norms etc. As we show, it thus makes little sense to understand mentality as spatially and temporally located (except in some trivial sense). Yet processes and states, on most understandings of these terms, are in fact spatially and temporally locatable items. Indeed, this is presumably why talk of processes and states is so common in the literature, since one can then aim to show that such processes and states either do or

Note that our wide view would be compatible with the question: where, for example, are the brain processes that realize or underpin a given mental episode? As we saw with our discussion of memory (see section 3.3), we allow that certain areas in the brain could indeed be part of the causal processes or mechanisms that underpin or realize, say, my ability to remember what I did last Monday. And such brain areas are of course localizable in space and time. Still, this would not support the further question: where in the brain is cognition realized? For if human mentality is not localizable (except in some trivial sense), then locating cognition in the brain (or anywhere else) is a non-starter.

However, some might challenge our wide view on the grounds that it is metaphysically suspect. Echoing the internalist worries mentioned in our introduction, they might object that if what makes a state or activity mental is simply that it sustains and explains intelligent behaviour, then all kinds of properties that play out against the personal and impersonal background would have to be considered. But then our wide view has the undesirable consequence that mentality spreads arbitrarily into the world and so fails to distinguish between those features of the world that obviously have mentality (people, animals) from those features of the world that do not (rocks, tables).

Yet going wide about mentality still allows for a clear distinction between those activities that are governed by the relevant practices and norms and those that are not. Think back to our writing example (section 3.2) and our wide view of cognition (section 3.4). Writing is a cognitive activity, not because of some property common to the underlying mechanisms that realize the activity, but in virtue of the fact that it sustains and explains why some behaviour is intelligent behaviour. The difference then between the actions of the first person (who is engaged in the activity of writing) and those of the second (who is simply making random marks on a page) is that only the actions of the first are recognisably mental, since only the actions of first can be understood to be

---

do not extend into the environment. Given our wide view however, we propose restricting talk of processes and states to the sub-personal level. One can still ask: do the causal processes and states that underpin a given cognitive capability extend into the environment? But of the two questions which active externalism raises e.g. Q1 and Q2, this question falls firmly within the remit of Q2. As such, it will not resolve the sorts of issues about mentality raised by Q1. And this then supports our point that if active externalism is understood as an attempt to answer Q1 via an appeal to Q2, then that debate requires dissolution, not solution.

intelligent behaviour. Contrarily, if it were to turn out that the actions of the second were *not* in fact random e.g. there was a pattern in the marks being made on the paper and that pattern had the discernible function to communicate, then we would also have grounds to call the actions of the second mental, since those actions could be understood as intelligent. Thus, a wide view of cognition retains a sense of what is and what is not mental. Indeed, by insisting on the independence between questions about what constitutes the mental (Q1) and questions about what is common to those causal processes or mechanisms that underpin or realize the mental (Q2), a wide view thereby clarifies why we only ascribe the term "mental" to some activities and not others. Hence, going wide neither implies that the mind spreads arbitrarily into the world nor that there is no distinction between those features of the world that have mentality from those that do not. For those features of the world that have mentality (people, animals) can be distinguished from those features of the world that do not (rocks, tables) because only the former and not the latter engage in recognisably intelligent behaviour.

Alternatively, others might object that going wide is scientifically problematic. It might be pointed out that since going wide now covers such a diverse array of behaviours and practices, then such an array is unlikely to display any underlying or low-level causal unities. Yet if the mark of a genuine scientific investigation is the identification of such unities, then a wide view is not scientifically tractable.

But this requires assuming that the only aim of science is the identification of underlying causal unities. We reject such a narrow view. As section 3.1 made clear, "relaxed naturalism" makes clear that going wide can be viewed as a naturalist position. Of course, understanding those features of the natural world that display an underlying causal unity from those features that don't will no doubt require employing different investigative strategies, methodologies and techniques. Yet this is to recognize the complexity of mentality. It is need not entail that mentality cannot be investigated scientifically.

To recap: we have argued that if active externalism is understood as an attempt to answer Q1 via an appeal to Q2, then the debate requires dissolution, not solution. However, what if the debate were instead understood as *only* an attempt to answer Q2? Could this offer a future for the debate?

Recall that Q2 concerns the property or properties common to the causal mechanisms that underpin or realise a given cognitive activity. As pointed out in section 2, one way to frame this question is in terms of cognitive systems. Indeed, this seems to be a favoured strategy in the literature (e.g. Rupert, 2010; Palermos, 2014; Pöyhönen, 2014). A cognitive system:

> "consists in the collection of mechanisms or capacities the contributions of which are highly correlated and which (unlike, say, air pressure) contribute distinctively to cognitive outcomes; this is the collection of capacities (or underlying physical mechanisms) each element of which contributes, as a member of overlapping subsets of mechanisms, to a wide range of cognitive outcomes. For convenience, think of the system as constituting a cognitive architecture." (Rupert, 2010, p12)

Note however that our wide view entails that the use of the term "cognitive" in the above definition can be nothing but derivative. For if mentality is truly wide, then there is nothing intrinsically "cognitive" about sub-personal mechanisms. The term "cognitive" only has application within personal level practices and activities. If we thus take seriously the above definition, that is, we understand a cognitive system as a collection of highly correlated mechanisms or capacities, then there literally are no such cognitive systems, since such systems will be bounded entities yet our wide view reveals cognition or mentality to be an unbounded, non-localisable phenomenon.

However, there could still be various sub-personal systems operative (in my head and/or spanning my brain and my body) whenever, say, I write or I remember. Such systems may in fact be part of the causal processes or mechanisms that underpin or realize my ability to write or my ability to remember. We can then ask: do the processes or states that make up such systems extend to include environmental objects or processes? And if they do, do such extended processes or states share a property or set of properties with those intracranial processes involved in the realisation of some

cognitive or mental capability? This question would be compatible with our wide view of mentality, since it would fall firmly within the remit of Q2. [17]

Q2 could possibly offer a future for the active externalism debate. Yet, as we made clear in section 2, Q2 is a question about explanatory practices in cognitive science. For determining whether or not there is a property or set of properties common to some sub-personal system will require first determining what is and what is not a genuine part of that system. And this will depend upon the phenomena we are trying to explain. For example, if are trying to explain REM sleep, it is not clear why we should include the local environment as part of the mechanistic base for that phenomenon. However, there likely will be grounds for viewing intracranial processes as genuine parts of the sub-personal system involved. In which case, if Q2 is to offer a future for the active externalism debate, then this will require assessing the explanatory (dis)advantages of active externalism over internalism.

Clark, for instance, has argued that active externalism offers the best explanation for some cognitive abilities. Conversely, Rupert has claimed that internal explanations win out. Yet, as Sprevak (2010) shows, inferences-to-the-best-explanation are not decisive in this back and forth, since such inferences cannot distinguish between active externalism and the Hypothesis of Embedded Cognition or HEMC, the idea that while non-biological objects or processes may be essential to the completion of some cognitive tasks, such objects or processes should not be regarded as part of the supervenience base for those tasks (Rupert, 2004).

[17] As suggested by an anonymous reviewer, one could read Clark as endorsing a cognitive systems view. As we point out in the text however, if one understands a "cognitive system" to be a collection of highly correlated mechanisms or capacities, then there are no such systems, according to our wide view. However, there is a way of reading Clark that would be compatible with our wide view. For, as we also acknowledge, there could be various sub-personal systems operative whenever you or I engage in a cognitive activity, like writing or remembering. If Clark was read as claiming that the processes or states that make up sub-personal systems extend to include environmental objects or processes and hence active externalism is true, then this would be compatible with our wide view, since this reading would fall within the remit of Q2. Moreover, Q2 is about explanatory practices in cognitive science. Active externalism, on this Clarkian reading, would then be the claim that externalism has an explanatory advantage over internalist positions. This reading would be compatible with our claim that a possible future for the active externalism debate could be one that centers on explanatory practices in cognitive science and so is only about Q2.

Moreover, even staunch defenders of internalism find nothing objectionable in HEMC. For example, Adams and Aizawa write:

> "Why not aim for a scientific and philosophical contribution that is empirically plausible and interesting? Sticking with the claim that cognitive processes are causally dependent on bodily and environmental processes is, of course, an orthodox view in early twenty-first century cognitive science. This is just the familiar rejection of Leibnizian monadology. That hypothesis alone is not an advance. But one might take positive steps forward in cognitive science by spelling out the kinds and scope of causal dependencies between cognition, body, and environment. Indeed, there is a much more conservative segment of the embodied and embedded cognitive science literature that does just this." (Adams & Aizawa 2010, p. 177)

That is, internalists see something favourable in recognising the important role environmental objects or processes play in cognition. In other words, as far as explanation is concerned, externalism has already has earned its explanatory spurs. Sprevak also makes a similar point:

> "The debate about the explanatory value of HEC [Hypothesis of Extended Cognition] to cognitive science is not about whether or not the mind extends. That issue is simply not sensitive to the explanatory practice of cognitive science. The debate could be about whether transcranial kinds should be allowed into cognitive science at all. Alternatively, it could be about whether the explanation of cognitive processes should be a purely internal matter (à la HINT) [Hypothesis of Internal Cognition]. But on both scores, it seems that the externalist has already won. Transcranial kinds are already doing useful work in psychology as the studies of Gray and Ballard show. And psychology no longer assumes that cognition can wholly be explained in the internalist way envisaged by HINT." (Sprevak, 2010, p18)

Hence, if the debate over active externalism is understood as nothing more than the claim that cognitive science can gain an explanatory edge by looking at extended

systems to *explain* cognitive abilities, then internalists and externalists are already singing from the same hymn sheet. There could still be a debate about whether a given cognitive ability is best explained by relying purely on internal mechanisms or by casting the explanatory net wider. This is indeed an important question, or rather a question that has to be answered separately for each and every cognitive ability. Poyhönen (2014), who defends an explanatory approach to the problem of demarcating cognitive systems, concludes as much:

> "I contend that choices between externalist and internalist classification strategies are necessarily more local, and based partly on the epistemic aims of the scientific field in question." (Pöyhönen 2014, p. 755)

But, as also acknowledged by Pöyhönen, this then turns the debate about active externalism into something very different. For example, it could be the case that for certain cognitive phenomena, latter-day internalists are correct in holding that internal mechanisms can do the required explanatory work, while for other cognitive phenomena, latter-day externalists are correct in that better explanations can be gained by including extended mechanisms (Clark could perhaps be read in this way – see footnote 18 – although he has not been read in this way by important contributors to the debate).

However, even if the latter-day internalist wins one such debate (but loses others), this will not be because of the sorts of arguments internalists like Adams and Aizawa have put forward. In other words, it will not be because internal mechanisms have some elevated metaphysical status, which renders such mechanisms "cognitive" or "mental". Rather, it will be because including extended mechanisms does not confer any explanatory benefits over and above those provided by internal mechanisms. In which case, while reconfiguring the debate to be one of explanation may certainly provide a future for active externalism, it also entails that the debate as understood by those on opposite sides of the internalist/externalist dividing line, that is, understood as an attempt to lay bare the essential core of cognition or answer Q1, is in fact all but over.

Bibliography

Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14, pp. 43-64,

Adams, F., & Aizawa, K. (2010a). Defending the bounds of cognition. In: R. Menary (ed.) *The Extended Mind*, Cambridge (MA): MIT Press, pp. 67-80

Adams, F., & Aizawa, K. (2010b). The value of cognitivism in thinking about extended cognition. *Phenomenology and the Cognitive Sciences*, 9(4), 579–603.

Adams, F., & Garrison, R. (2013). The Mark of the Cognitive. *Minds and Machines*, 23(3), 339–352.

Bennett, M. and Hacker, P., "The conceptual presuppositions of cognitive neuroscience", in: D. Robinson (ed.) *Neuroscience and Philosophy: Brain, Mind, and Language,* p. 127 – 162, New York: Columbia University Press, 2007.

Clark, A. (2010). Coupling, Constitution, and the Cognitive Kind: A reply to Adams and Aizawa. In: R. Menary (ed.) *The Extended Mind*, Cambridge (MA): MIT Press, pp. 81-100

Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58, pp. 7-19.

Hatfield G. (2014). Cognition. In: L. Shapiro (ed.) *The Routledge Handbook of Embodied Cognition*, London: Routledge, pp. 361-373

Hutto, D. & Myin, E. (2013). *Radicalizing Enactivism Basic Minds without Content*. Cambridge (MA): MIT Press

Hutto, D. & Satne, G. (2015). The natural origins of content. *Philosophia*, 43, pp. 521–536.

Menary, R. (2010). *The Extended Mind.* Cambridge (MA): MIT Press.

Moyal-Sharrock, D. (2013). Wittgenstein's razor: the cutting edge of enactivism. *American Philosophical Quarterly*, 50, pp. 263-279.

Myin, E. & Zahidi, K. (2012). Uitgebreid, complementair of omvattend? Het waar en hoe van het mentale. *Algemeen Nederlands tijdschrift voor wijsbegeerte*, 104, p. 173-188

Myin, E. & Zahidi, K. (2014). The extent of memory: from extended to extensive mind. In: D. Moyal-Sharrock et al.(Eds.), *Mind, language and action: proceedings of the 36th International Wittgenstein Symposium*. Berlin: De Gruyter.

Palermos, S, O. (2014). Loops, constitution and cognitive extension. *Cognitive Systems Research*, 27, p. 25-41.

Pöyhönen, S. (2014). Explanatory power of extended cognition. *Philosophical Psychology*, 27(5), 735–759.

Rowlands, M. (2010). *The New Science of the Mind. From Extended Mind to Embodied Phenomenology*. Cambridge (MA): MIT Press

Rupert, R. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy*, 101, pp. 389–428.

Rupert, R. (2010). *Cognitive Systems and the Extended Mind*. Oxford: Oxford University Press.

Sprevak, M. (2010). Inference to the hypothesis of extended cognition. *Studies in History and Philosophy of Science Part A*, 41(4), pp. 353–362.

Thompson, E, Stapleton, M. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi*, 28 (1), p. 23-30.

Vilar, P. (1984). *A History of Gold and Money (1450-1920)*. London: Verso.

Wheeler, M. (2005). *Reconstructing the Cognitive World*. Cambridge (MA): MIT Press.