ACTIVE EXTERNALISM AND EPISTEMIC INTERNALISM

J. Adam Carter & S. Orestis Palermos
University of Edinburgh

Abstract. Internalist approaches to epistemic justification are, though controversial, considered a live option in contemporary epistemology. Accordingly, if 'active' externalist approaches in the philosophy of mind—e.g. the extended cognition and extended mind theses—are in principle incompatible with internalist approaches to justification in epistemology, then this will be an epistemological strike against, at least the prima facie appeal of, active externalism. It is shown here however that, contrary to pretheoretical intuitions, neither the extended cognition nor the extended mind theses are in principle incompatible with two prominent versions of epistemic internalism—viz., accessibilism and mentalism. In fact, one possible diagnosis is that pretheoretical intuitions regarding the incompatibility of active externalism with epistemic internalism are symptomatic of a tacit yet incorrect identification of epistemic internalism with epistemic individualism. Thus, active externalism is not in principle incompatible with epistemic internalism per se and does not (despite initial appearances to the contrary) significantly restrict one's options in epistemology.

Keywords: extended cognition, extended mind, epistemic internalism

1 Two Questions of Incompatibility

Ever since the Putnam/Burge-style thought experiments led to the institution of content externalism—viz., the thesis that mental contents are widely individuated—as orthodoxy in mainstream philosophy of mind, there has been a lively research program centred around the following question: how, if at all, can externally individuated mental contents be reconciled with the independently plausible epistemological insight that we have privileged access to the contents of our own minds?

The ensuing 'externalism and self-knowledge' literature1 naturally gave rise to a more general kind of philosophical question: can content externalism in the philosophy of mind be reconciled with internalism in epistemology2—the thesis that (put simply) epistemic justification is a 'purely internal matter'3? If the answer is no, then the overwhelming plausibility of externalism about mental contents would seem to entail externalism more broadly in the theory of knowledge4.

---

2 See fn. 4.
3 See here Bach (1985, 250). This is of course a rough characterisation of the position. In what follows we will be addressing epistemic internalism under two specific forms: an accessibilist form and a mentalist form. For a helpful overview, see Pappas (2005).
4 Bonjour (1992, 136), in an oft-cited passage that characterizes one line of motivation for incompatibilism, remarks that: "The adoption of an externalist account of mental content would seem to support an externalist account of justification in the following way: if part or all of the content of a belief is inaccessible to the believer, then both the justifying status of other beliefs in relation to that content and the status of that content as justifying further
Clark & Chalmers (1998) refer to content externalism, as it features in these debates as ‘passive’ to distinguish this kind of externalist position from the more metaphysically robust externalist theses that fall under the heading of ‘active externalism’—and of which two notable forms in the recent literature include the hypothesis of extended cognition, and the extended mind thesis. Put roughly, the former thesis insists that cognitive processes can include parts of the world external to the biological agent. The comparatively more radical extended mind thesis maintains that mental states—such as beliefs and desires—can be realised in part by things external to the biological agent.

We therefore see that regardless of whether content (passive) externalism entails epistemic externalism, there is a further question we can ask: In what ways, if any, are active externalism and epistemic externalism entangled? Does active externalism entail epistemic externalism? More specifically: are either the extended cognition thesis or the extended mind thesis compatible with epistemic internalism?

Even though, as indicated above, there is an ongoing, live debate with respect to the compatibility of content (passive) externalism with epistemic internalism, there is virtually no discussion with respect to the issue of whether active externalism is compatible with epistemic internalism. An unsurprising disparity, of course, given that passive externalism has been around for almost 40 years now, but not so for active externalism, which was first introduced in the turn of our century; active externalism is still very much under development and its input to epistemology in general has only just started receiving attention—let alone its compatibility with specific internalist views. Nevertheless, given epistemic internalism’s traditional emphasis on the justificatory importance of what lies underneath the organism of the epistemic agent and, in its most common versions, also on what is within reach to his conscious psychology, a natural pre-theoretical thought will be to insist that, regardless of whether content (passive) externalism can be reconciled with epistemic internalism, the more radical thesis of active externalism surely cannot.

In the next section, we’ll sharpen this idea, which for the present purposes we’ll call the incompatibility thesis. But before doing so, it is helpful to locate the issue within a wider

---


7 The reason why we take the last two questions to be identical is because epistemic externalism is standarly defined as the denial of epistemic internalism (however the latter is defined).

8 Cf. Smithies (Forthcoming), for instance, who suggests that active externalism is incompatible with mentalist epistemic internalism. Though, again, this is by and large unexplored territory.
perspective: if the incompatibility thesis is true, then this is a *prima facie* strike against active externalism.

2 The Incompatibility Thesis: Four Motivating Claims

The incompatibility thesis draws traction from four separate *prima facie* plausible 'sub'-theses, which arise once we consider whether either extended cognition or the extended mind thesis could be reconciled with *either* of the two most prominent strands of epistemic internalism: (i) *Accessibilist* justification internalism (hereafter, *accessibilist J*-internalism) and *mentalist* justification internalism (hereafter, *mentalist J*-internalism).

According to *accessibilist J*-internalism, an agent can determine by reflection alone the factors that would make her beliefs epistemically justified; an associated 'negative' commitment of the accessibilist position is that agents do not diverge in the extent to which their beliefs are justified provided that they do not diverge in what is accessible to them by *reflection alone*. Of course, reflection is here closely associated with consciousness, which, in turn, is presupposed, to reside within the agent’s head, in accordance with traditional wisdom.

Mentalist J-internalism, by contrast, is the thesis that, as Conee and Feldman (2001) put it, ‘a person’s beliefs are justified only by things that are internal to the person’s mental life’; correlative, what justifies one’s beliefs supervenes entirely on one’s mental states such that, individuals who do not differ in their mental states do not diverge in the extent to which they are epistemically justified in their beliefs. And, of course, given received thinking about the supervenience base of mental states as seated within the neurobiology of the agent, mentalist J-internalists will be inclined to follow Chisholm’s thinking of “justifiedness as a purely internal matter” in a way that is not ‘external’ in the sense that active externalists take cognition to be.

In light of the distinction between mentalist and accessibilist J-internalism, we can draw out four separate claims that, taken together, would establish the incompatibility of active externalism and epistemic internalism:

(i) If extended cognition is true, accessibilist J-internalism is false.
(ii) If extended cognition is true, mentalist J-internalism is false.
(iii) If the extended mind thesis is true, accessibilist J-internalism is false.
(iv) If the extended mind thesis is true, mentalist J-internalism is false.

---

9 Although, of course, depending on one’s starting point, the exact opposite can also be said: if the incompatibility thesis is true, then, to the extent that active externalism is a live option in the metaphysics of mind, this would constitute a *prima facie* strike against epistemic internalism.

10 See here, for instance Chisholm (1977) and Bonjour (1985, ch. 2) for classic defences of the accessibilist position; for a recent overview, see also Madison (2010).

11 See, for instance, Bergmann (2006).

12 As Carl Ginet (1975) puts the idea: “Every one of every set of facts about S’s position that minimally suffices to make S, at a given time, justified in being confident that p must be directly recognizable to S at that time. By ‘directly recognizable’, I mean this: if a certain fact obtains, then it is directly recognizable to S if and only if, provided that S at that time has the concept of that sort of fact, S needs at that time only to reflect clear-headedly on the question of whether or not that fact obtains in order to know that is does obtain” (34).
The main objective of this paper will be to show that (despite their initial plausibility) it’s not clear whether any of (i)-(v) should be accepted; the most natural arguments for these four theses come up short, and therefore—perhaps surprisingly—there is at the end of the day no good reason to think that active externalism is, in principle, incompatible with epistemic internalism.

Of course, the above may be an understatement, the reason being that some of the moves required to block the arguments for the incompatibility of active externalism and epistemic internalism will alter what the latter position has traditionally been thought to amount to beyond recognition. For instance, as we shall see, some of the moves for retaining active externalism while also adhering to the letter of epistemic internalism in both its accessibilist and mentalist guises, will insist that reflection can take place outside the agent’s head or that an agent’s justificatory mental states extend beyond his organism, respectively. Of course, these are rather odd moves to make—though perhaps not to an equal extent—and it is far from clear what’s left of epistemic internalism once either of them is made. As we shall, argue, however, none of these moves actually threatens the spirit of epistemic internalism per se but only the spirit of what has lately come to be known in the literature as epistemic individualism (e.g. Pritchard & Kallestrup, 2013). As far as we can see, however, there is no reason why we should run these two positions together as it has implicitly been the case so far. As we shall, argue, however, none of these moves actually threatens the spirit of epistemic internalism per se but only the spirit of what has lately come to be known in the literature as epistemic individualism (e.g. Pritchard & Kallestrup, 2013). As far as we can see, however, there is no reason why we should run these two positions together as it has implicitly been the case so far.

This is the uncharted dialectical terrain we here wish to explore in the hope to close the gap in the literature with respect to the possible ways, if any, active externalism can combine with epistemic internalism. The plan of the paper is as follows: §3 critically assesses claims (i)-(ii), and section §4 claims (iii)-(iv). We conclude with some general comments on why the incompatibility thesis may initially (but not eventually) seem true and what pursuing (some forms of) epistemic internalism along the lines suggested by (some forms of) active externalism might look like.

3 Extended Cognition and Epistemic Internalism

3.1 Extended cognition

One simple way to make sense of the idea of extended cognition is via common-sense functionalism; this is the primary rationale driving Andy Clark and David Chalmers’s (1998) parity principle, a principle which tells us, by analogical reasoning, under what conditions something should be ‘ruled-in’ as part of a cognitive process:

---

13 That is, there is no reason to suppose that epistemic internalism entails epistemic individualism, even though this is tacitly taken to be the case.

14 The parity principle should be understood as an intuition pump for assessing whether a part of the world should be included as part of a cognitive process. It is important to distinguish this principle telling us what is ruled-in and ruled-out, from what would be conditions specifying in virtue of what something is a cognitive process. The issue of in virtue of what may something count as part of a cognitive process is the issue of metaphysical grounding, which is, as Fine (2012, 37) puts it, ‘a distinctive kind of metaphysical explanation, in which explanans and explanandum are connected, not through some sort of causal mechanism, but through some constitutive form of determination.’ Unlike the extended mind thesis, which relies almost entirely on common-sense functionalism for its philosophical motivations, extended cognition proponents have reached beyond common-sense functionalism to dynamical systems theory (e.g. Chemero (2009), Froese et al. (2013), Sutton et al. (2008), Theiner et al. (2010),
Parity Principle: If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is part of the cognitive process. (Clark & Chalmers 1998, 8)

The parity principle is aimed at preserving a kind of egalitarian spirit about the cognitive; in particular, it forbids recourse to the idea that ‘it is just on account of the fact that something is external to the skin and skull’ that it is excluded as part of a cognitive process.

To see how the parity principle ‘rules-in’ a part of the world as part of a cognitive process in a concrete case, it is helpful to consider here Clark and Chalmers’s often referenced case of ‘Otto’:

Otto: Otto suffers from Alzheimer’s disease, and like many Alzheimer patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory.

The case is meant to elicit an appreciation for how an individual can rely on an external object (e.g. a notebook) in such a way that the external object is playing the same functional role, vis-a-vis storage and retrieval of information, as an ordinary agent’s biological memory. Accordingly, Clark & Chalmers think there is no good reason not to view Otto’s cognitive process as a kind of ‘extended memory’, where the extended memorial process is partly realized by something extra-cranial. Moreover, in order to make this common-sense functionalist claim more plausible, Clark (2010) notes that, just as in the case of biological memory, the availability and portability of the external resource of information should be crucial. Accordingly, he offers the following set of (‘glue and trust’) criteria to be met by non-biological candidates for inclusion into an individual’s mind:

1) “That the resource be reliably available and typically invoked”.
2) “That any information thus retrieved be more-or-less automatically endorsed. It should not usually be subject to critical scrutiny.
3) “That information contained in the resource should be easily accessible as and when required.” (Clark 2010, 46)

3.2 Extended cognition and accessibilist J-internalism

Let’s now assess claim (i) of the incompatibility thesis:

Tollefsen & Dale (2011), Palermos (2014) to offer a kind of metaphysical explanation suited to answer the question in virtue of what is something a cognitive process.

15 While this paper was first published in 2010, it has been available online since 2006. The ‘glue and trust’ criteria, however, had already made their appearance in Clark & Chalmers (1998), although the phrasing was somewhat different.
(i) If extended cognition is true, accessibilist J-internalism is false.

Put in very simple terms, it is tempting to think (i) is true because it’s natural to assume something like the following: ‘Extending’ cognitive processes is going to have a knock-on effect of making certain facts relevant to justifying one’s beliefs fall outside the realm of what is accessible to the agent by reflection alone, in a sense the accessibilist J-internalist will find particularly objectionable.

Toward this end, let’s consider an example aimed at showing that, if extended cognition is true, then the following is also true: contrary to what the accessibilist J-internalist must say, even though A and B do not diverge with respect to what is accessible by reflection alone (as is relevant to \( p \)), A and B diverge vis-à-vis their epistemic justification for \( p \). If a case could show this much, then we can envision a straightforward argument for accepting (i).

A case with this kind of structure won’t be hard to generate: let’s simply borrow Clark & Chalmers’ hero, Otto (as described in §3.1), who has relied on his notebook for years such that it is now integrated into his cognitive system, and for whom the notebook was said to play the functional role of biological memory (by reference to the parity principle and the ‘glue and trust’ criteria). Now, let Otto* be like Otto, except that Otto* has only just begun using the notebook to compensate for his failing biological memory—and as such, Otto*’s notebook is not playing the same role (i.e., that of biological memory) that Otto’s notebook plays for him (the easiest way to explain the intuitive disparity between the two agents is by reference to the ‘glue and trust’ criteria).

Now, stipulate that Otto and Otto* both have written down the time of their doctor’s appointment (noon) in their respective notebooks—a proposition we can represent as ‘\( p \)’. Moreover, let T be a time right before Otto and Otto* consult their notebooks to retrieve the information. With that in mind, an argument for (i) might unfold as follows:

1. Otto and Otto* do not diverge vis-à-vis what is accessible to them by reflection alone, at T.
2. If extended cognition is true, then, intuitively—that is with no specific theory of justification in mind—at T, the entry in Otto’s notebook is epistemically justified for Otto to a greater extent than this entry is for Otto*.
3. If accessibilist J-internalism is true, then if Otto and Otto* do not diverge in what is accessible to them by reflection alone at T, then they do not diverge in the extent to which they are justified vis-à-vis a given proposition \( p \), at T.
4. Therefore, if extended cognition is true, accessibilist J-internalism is false.

Premise (3) is true by definition; premise (1) is supported by the thought that Otto and Otto* are obviously in the same position as concerns what is reflectively accessible to them, as this

---

16 The most natural kind of case to devise to this end will be one that tries to exploit the thought that two agents could be relevantly the same vis-à-vis what is accessible to them by reflection alone, while letting just one of them enjoy what, by reference to the parity principle, constitutes an extended cognitive process, and in such a way that must consequently break the symmetry of their epistemic positions.

17 For recent discussions of cognitive integration within the context of extended cognitive processes, see here Pritchard (2010) and Palermos (2011; 2014).
bears on \( p \). To be clear, Otto and Otto* may indeed differ on what is immediately accessible to them (due to the fact that they do not satisfy the ‘glue and trust’ criteria to the same extent), but given that they both need to have perceptual access to their notebooks, what is accessible to them via reflection alone has to be the same. The premise doing important work is (2); (2) is motivated by the thought that the difference between Otto and Otto*—namely, that (on the extended cognition proposal) the former enjoys an extended cognitive process the latter lacks—will be a kind of epistemic difference maker with respect to Otto and Otto*’s respective justification vis-à-vis \( p \).

But an argument along these lines runs into trouble with both premises (1) and (2). Firstly, the issue with premise (1) concerns reflection: either the notion of reflection operative in Otto’s extended memorial process is to be conceived of as non-extended (e.g. as a process realised entirely intracranially), or as extended. Interestingly, there’s some difficulty in maintaining the former position without begging the question against the proponent of extended cognition. For example, one may argue that reflection should be here understood as identical to introspection and that the obviously perceptual access of Otto to the relevant information precludes it from being ‘reflectively’ accessible. The problem with this line of thought, however, is that, as far as the common-sense functionalist is concerned, introspection is used to refer to the process of looking within one’s own cognitive system. Accordingly, since Otto’s cognitive system extends to his notebook, perceiving the relevant information with his eyes would qualify as a case of introspection. Therefore, it would seem question-begging to deny that extended memorial process can be reconstrued as involving the extended functional analogue of reflection. But of course, as soon as we think of extended memory as involving anything like extended reflection, Premise (1) comes out false.

Suppose though that we simply grant premise (1) for the sake of argument. There remains a further and more serious problem with Premise (2). For one thing, at \( T \), it could be pointed out that neither Otto* nor Otto seems to count as actually believing the target proposition. After all, at \( T \), they both have yet to look \( p \) up. Now this observation would be neither here nor there if the extended cognition thesis simply entailed that Otto can enjoy extended mental states such as extended dispositional beliefs at \( T \) (i.e., count as holding ‘beliefs’ that are inscribed in his notebook but hasn’t looked up yet). But strictly speaking the hypothesis of extended cognition does not entail the extended mind thesis and so doesn’t have this strict consequence (we will return to this in §4).

---

18 As we’ll see later in this section, this assumption may be controversial by the lights of a proponent of extended cognition. Nonetheless, to whatever extent an argument for the incompatibility of extended cognition and accessibilist J-internalsim will get off the ground, something like Premise (1) will be taken as evident.

19 Consider that the proponent of extended cognition (who takes Otto’s process to be an extended memorial process) might insist that the kind of common-sense functionalism that rules Otto’s memorial process to be extended will also entail that to whatever extent reflection is integral to memory, there must be an extended way of conceiving of this. To rule this out ex ante is to thus begin by rejecting a philosophical principle of extended cognition. Notice further that this worry about premise 1 seems to show its head no matter how we characterize Accessibilist J-internalsim; it wouldn’t matter, for example, if we replaced ‘by reflection alone’ with Greco’s (2005) ‘from the agent’s internal perspective’ or with a related ‘awareness’ condition.

20 Just consider that, if Otto, at \( T \), were to form an occurrent belief that \( p \), at \( T \), it would not be justified for Otto, at \( T \). Indeed, this would be the analogue (in the non-extended case) of forming a belief about something that happened in one’s past prior to consulting one’s biological memory.
But then, if neither Otto* nor Otto count as believing \( p \) at \( T \), it’s hard to see how the extent to which they would be justified, were this extent to differ, would be a difference of doxastic justification at \( T \). It is natural to suppose though that the relevant difference in justification is a difference in dispositional justification. As Schwitzgebel (2014) puts it, ‘A subject dispositionally believes \( P \) if a representation with the content \( P \) is stored in his memory or “belief box”’\(^{21}\). Assume for the sake of argument that extended cognition implies that Otto has a dispositional belief that \( p \), even though Otto* does not\(^{22}\). Even if this were so it’s not clear that premise (2) will be established. After all, plausibly, \( p \) is dispositionally justified for \( S \) (on an accessibilist J-internalist proposal) just in case the following counterfactual is true: were \( S \) to contemplate \( p \), \( S \) would be able to determine by reflection alone the factors that justify \( p \)^{23}. But it’s not evident how this counterfactual holds for Otto in a way it would not for Otto*\(^{24}\).

Rather than doxastic or dispositional justification, then, it would seem like any difference in justification at all implied by extended cognition would be propositional justification; the operative claim would be that despite not differing in what is reflectively accessible to Otto and Otto*, \( p \) is propositionally justified for Otto to a greater extent than for Otto*. But does extended cognition really imply that Otto and Otto* diverge vis-à-vis propositional justification (with respect to \( p \)) at \( T \)?

A familiar line on the difference between propositional justification and doxastic justification goes like this: one is propositionally justified in believing \( p \) iff one’s evidence sufficiently supports \( p \)^{25}, whereas doxastic justification involves actually basing one’s belief on the evidence that makes one’s belief propositionally justified\(^{26}\). But there is no obvious reason to think that viewing Otto’s process as extended entails that his evidence is thereby different than the evidence available to Otto*, at \( T \)^{27}.

\( \text{21} \) Schwitzgebel (2014, §2.1).

\( \text{22} \) Although we should note, again, that the hypothesis of extended cognition (as opposed to the extended mind thesis) does not entail the existence of extended dispositional beliefs.


\( \text{24} \) Unless, of course, just as we argued above, we allow for reflection to be extended in the Otto but not in the Otto* case. What this means, however, is that premises (1) and (2) cannot be true at the same time.

\( \text{25} \) For a representative statement of this idea, see Neta (2007) and Firth (1978).

\( \text{26} \) As Kvanvig puts it, ‘Doxastic justification is what you get when you believe something for which you have propositional justification, and you base your belief on that which propositionally justifies it.’ See also here Pollock & Cruz (1999) and Swain. Cf. Turri (2010) who reverses the order of explanation and defines propositional justification in terms of doxastic justification.

\( \text{27} \) In other words, the matter of whether an agent satisfies conditions similar to the ‘glue and trust’ criteria is plausibly going to be orthogonal to whether a proposition \( p \) or a sense datum counts as evidence for that agent or not. But even if something like this could be established such that a proponent of the incompatibility of extended cognition and accessibilist J-internalism were able to demonstrate that extended cognition implies that for two subjects who don’t differ vis-à-vis what is reflectively accessible to them, they nonetheless differ vis-à-vis propositional justification, this would be a kind of Pyrrhic victory. Michael Bergman (2007, 19) surely speaks for many in his remark that ‘I’ve never been much interested in propositional justification. Insofar as I’ve been interested in epistemic justification, it’s doxastic justification that I care about.’ This is of course not to say that propositional justification is not philosophically important at all; but rather, that the kind of epistemic justification of primary interest to epistemologists is doxastic, not propositional, justification. Accordingly, establishing incompatibility only at the propositional level will be viewed as coming up importantly short.
3.3 Extended cognition and mentalist J-internalism

It’s not at all clear, then, that extended cognition is in principle incompatible with accessibilist J-internalism. Let’s turn now to considering whether extended cognition faces more difficult compatibility problems if we think of epistemic internalism along mentalist rather than accessibilist lines. In other words, we’ll be evaluating what we called claim (ii) in §2:

(ii) If extended cognition is true, mentalist J-internalism is false

Recall that the mentalist J-internalism thesis states that a person’s beliefs are justified only by things that are internal to the person’s mental life. A typical understanding of this thesis is as a supervenience claim to the effect that epistemic justification for S, at T, supervenes on S’s mental states at T. A corollary is that S and S* do not differ in the extent to which they are justified at T provided they do not differ vis-à-vis their mental states at T.

The argument for the incompatibility of extended cognition and mentalist J-internalism stated in (ii) will take as a natural starting point the claim that:

(1) Otto and Otto* are in the same mental state at T (just prior to consulting the notebook).

And then, the argument runs as follows:

(2) If extended cognition is true, then intuitively—that is with no specific theory of justification in mind—Otto and Otto* diverge in the extent to which they are justified in p at T.
(3) If mentalist-J-internalism is true, Otto and Otto* do not diverge in the extent to which they are justified in p at t. (From Mentalist J-internalism and (1)).
(4) Therefore, if extended cognition is true, mentalist J-internalism is false. (From (2) and (3)).

Again (3) is true by definition, and the controversial premises will be (2) and perhaps (1). We say (1) might be controversial because, even though the extended cognition hypothesis does not entail and neither does it motivate in any obvious way the existence of extended mental states, the existence of such states would still be argued for on the basis of the extended cognition hypothesis by following some form of reasoning along the following lines: just as biological memorial processes are thought to support dispositional beliefs, extended memorial processes can be thought to support extended dispositional beliefs. Accordingly, to the extent that this line of thinking is right, there is a case for rejecting (1) even if one denies the stronger extended mind thesis.

However, even if (1) is granted, there is cause for rejecting (2), that parallels the kind of discussion vis-à-vis accessibilist J-internalism. Specifically, the point here is that if we grant (1) is true and that Otto and Otto* are in the same mental states at T, then even if the extended cognition thesis is true and thus Otto’s notebook (but not Otto*’s) counts as part of his memorial process, it’s hardly clear that Otto and Otto* would thereby differ vis-à-vis their justification for p at T. The line of thought in the case of mentalist J-internalism parallels the discussion in §3.2: the difference in justification implied by extended cognition and needed to
make the argument for (ii) go through will not be doxastic, nor dispositional, nor for that matter propositional. Premise (2) ends up accordingly at best undermotivated.

The discussion in §2 (both §2.1 and §2.2) thus shows that, contrary to what might have been suspected, the extended cognition thesis does not obviously stand at odds with epistemic internalism (of either the accessibilist or mentalist varieties).

4 Epistemic Internalism and the Extended Mind Thesis

4.1 Accessibilist J-internalism and the Extended Mind Thesis

Claims (i) and (ii) supporting the incompatibilist thesis did not turn out promising. Both claims however concerned the extended cognition brand of active externalism. Let’s now turn to the issue of whether the extended mind thesis entails epistemic externalism. In particular, let’s begin by assessing the accessibilist leg of this aspect of the incompatibility thesis, by turning our attention to claim (iii).

(iii) If the extended mind thesis is true, accessibilist J-internalism is false.

The case of Otto and Otto*, even if not ultimately successful in establishing the incompatibility of extended cognition and epistemic internalism, looks, at least prima facie, like exactly the sort of case that might establish the incompatibility of the extended mind thesis and epistemic internalism.

A first way to make the argument (vis-à-vis the accessibilist J-internalism) will be to begin with a premise to the effect that Otto and Otto* (from our case in §3) do not diverge vis-à-vis what is accessible to them by reflection alone (and then to show that, an implication of the extended mind thesis is that despite being on a par vis-à-vis what is reflectively accessible, they will nonetheless diverge vis-à-vis their epistemic justification\(^{28}\)). The starting premise is this:

\[(1) \text{ As regards what bears directly on the matter of whether } p \text{ (as stipulated in the Otto/Otto* case) is justified, what is accessible to Otto and Otto* by reflection alone is the same, at } T.\]

Of course, even the stipulation in (1) is at least somewhat controversial for reasons we’ve already considered in §3.2. But as we’ve said, something like (1) is needed in order to go on to get the incompatibility argument off the ground.\(^{29}\) We’ll consider a variant on (1) shortly, but for now, with (1) in hand, the rest of the reasoning proceeds as follows:

\(^{28}\) Note that the structure of this argument bears some expected structural similarities vis-à-vis the argument considered on behalf of the proponents of the incompatibility of accessibilist J-internalism and the hypothesis of extended cognition; in each case, we find a very natural way to move from one of the active externalist theses to incompatibilism; importantly, of course, as the hypothesis of extended cognition does not entail the more radical EMT, the key premises are substantially different.

\(^{29}\) But to make more explicit the reason something like (1) is needed is that, without insisting on a premise like (1), then one isn’t in a position to demonstrate that the antecedent in the conditional in (iii) can be true while the consequent false. Because the consequent itself involves a conditional, the claim in (iii) is of the form:
(2) If (1), then by accessibilist-J internalism, Otto and Otto* do not differ in the extent to which they are justified vis-à-vis \( p \).

And then:

(3) If the extended mind thesis is true, then, intuitively—that is with no specific theory of justification in mind—Otto and Otto* differ in the extent to which they are justified vis-à-vis \( p \), at \( T \).

The motivation for premise (3) is the following: given that Otto’s notebook is (by parity) playing the role of Otto’s biological memory, but the notebook is not playing this role for Otto*, Otto has (by the extended mind thesis) an extended dispositional belief that \( p \) at \( T \), that Otto* lacks, at \( T^{30} \). Now, in the absence of a compelling way to account for how dispositional beliefs will be justified\(^{31} \), or at least of a way for showing how we can solve some of the problems facing the justificational status of dispositional beliefs (see also below), then (as the argument continues):

(4) The extended mind proponent must deny a claim that the proponent of accessibilist J-internalism must accept: that Otto and Otto* do not differ in the extent to which they are justified vis-à-vis \( p \), at \( T \).

(5) Therefore, if the extended mind thesis is true, accessibilist J-internalism is false.

While this isn’t the only way to argue for (iii), it is a natural first pass for establishing that the countenancing of extended mental states will have as a consequence a denial of what the accessibilist J-internalist is committed to—viz., in short, that sameness in reflective accessibility entails sameness in justification.

Is it compelling? There are several lines of resistance available. An obvious first line will simply insist (\( \text{per } \$3.2 \)) that premise (1) will come out true only on a conception of reflection that the proponent of extended cognition will find objectionable (and thus what follows won’t be an implication of accepting the extended mind thesis). But we won’t rehearse this line again.

Rather, consider here that a proponent of the incompatibility conclusion could try to circumvent entirely some of the \( \$3.2 \)-style objections to (1) by simply re-casting the argument in

\[ [(A \rightarrow (B \rightarrow C))] \]. The counterexample thus will involve supposing that \( A \) and \( B \) can be true while \( C \) false, and thus will be a case where we stipulate that the extended mind thesis is true and that two agents don’t diverge vis-à-vis what is reflectively accessible to them, but as an implication of extended mind thesis do diverge in their respective justification. Such a case will thus need to be one where it is insisted that an agent is going to count the extended mind thesis as having an extended mental state while also not diverging in what is reflectively accessible to her, vis-à-vis a counterpart without an extended state. Accordingly, unless something like one is granted, it’s not clear at all how one would be in a position to demonstrate the falsity of (iii) via counterexample.

\(^{30}\) Note that Clark & Chalmers (1998) take it that the original case of Otto suffices to motivate by common-sense functionalist rationale both the extended cognition thesis and the extended mind thesis.

\(^{31}\) This unstable pairing would also violate what has been called the epistemic parity principle; see Carter (2013), Carter & Pritchard (2014) and Carter & Kallestrup (2014).
terms of a 'modified accessibility premise' in (1), styled after Kent Bach’s weaker accessibilist condition. The tweaked argument takes as the first premise a less controversial opening line:

(1*) As regards what bears directly on the matter of whether ‘p’ (as stipulated in the Otto/Otto* case), Otto and Otto* do not diverge in the extent to which they are aware or at least immediately capable of being aware of what makes p justified.

(1*) is much weaker than (1) and so harder to reject\(^\text{32}\). However, (1*) carries with it a need to tweak premise (2) as follows:

(2*): If (1*), then by (weak)-accessibilist-J internalism, Otto and Otto* do not differ in the extent to which they are justified vis-à-vis p.

And then the rest of the argument runs from there accordingly:

(3*) If the extended mind thesis is true, then, intuitively, Otto and Otto* differ in the extent to which they are justified vis-à-vis p, at T.
(4*) Therefore, the extended mind proponent must deny a claim that the proponent of accessibilist J-internalism must accept: that Otto and Otto* do not differ in the extent to which they are justified vis-à-vis p, at T.
(5*) Therefore, if the extended mind thesis is true, accessibilist J-internalism is false.

Interestingly, though, the proponent of the extended mind thesis is in a position to reject even the weaker (1*) on the grounds that—given that the notebook is cognitively integrated for Otto but not for Otto*—the entry in the notebook is (despite initial appearances) immediately accessible for Otto in a way that does in fact differ from the extent to which it is immediately available for Otto*.

On this line of thought, the extended mind proponent can draw an analogy to a non-extended case-pair and liken Otto to someone for whom multiplication tables are second nature, in comparison with someone who must ‘work it out’ mentally. Just as the answer to \[7 \times 9\] is available (in an inclusive sense) to both individuals, there is a clear sense in which the answer is immediately available to the former, for whom the process is second nature.

What goes for the mathematician goes for Otto, in that (given Otto’s integration of his notebook) the consulting of the notebook is second nature and immediate in a way that this is not for Otto\(^\text{33}\). But in so far as this is right, we have reason to reject the weaker (1*). Now, one may here object that the idea of ‘second nature’ it too vague to explain away the difference in

---

\(^{32}\) After all, premise (1*), unlike (1), is disjunctive; it comes out true so long as Otto and Otto* do not diverge either in the relevant awareness or capability of being immediately aware—whereas, the original (1) would be true only if they did not diverge vis-a-vis awareness. Thanks to an anonymous referee at Erkenntnis for encouraging clarification on this point.

\(^{33}\) For more discussion of the sense in which Otto’s process of consulting the notebook is second nature, and best not described as a series of discrete processes, see Clark (2010) and also Pritchard (2010) and Carter & Pritchard (2014).
the epistemic standing of the two agents. At this point, however, we can simply re-invoke the ‘glue and trust’ criteria who were proposed as a means for intuitively figuring out what may count as part of one’s mind and thereby of one’s self.

Still, however, even if (1*) were granted for the sake of argument, there remains another avenue for resisting the incompatibilist conclusion: by rejecting (3*). Consider that even though an implication of the extended mind thesis is that Otto* counts as having an extended belief at T that Otto+ lacks at T, an explanation for how Otto’s justification for p, at T, differs from the justification Otto+ has for p, at T, will have to be an explanation made in terms of doxastic justification. Notice, however, that if the extended belief enjoys justification at T, it’s at least not clear how the justification will be doxastic for the reason that it is controversial whether Otto bases his belief on what propositionally justifies the belief, at T34. Unless it can be shown how he does35, the epistemic justification will have to be explained in terms of the extended belief vis-à-vis dispositional belief, or alternatively, the difference in justification must be articulated in terms of a difference in propositional justification.

As noted before, however, going the dispositional route can be problematic (see pp. 7-8, above). The relevant counterfactuals that must be satisfied for Otto to be dispositionally justified (on an accessibilist J-internalist proposal) seem equally satisfied by Otto+. And the problem for offering such an explanation in terms of propositional justification applies mutatis mutandis from §3.2.

4.2 Mentalist J-Internalism and the Extended Mind Thesis

Perhaps the most prima facie unharmonious way of combining active externalism and epistemic internalism is by mixing the extended mind thesis with mentalist J-internalism. This speaks in favour of the initial plausibility of the fourth ‘sub’-thesis outlined in §2 as (iv), the claim that:

(iv) If the extended mind thesis is true, mentalist J-internalism is false.

Mentalist J-internalists maintain a supervenience thesis to the effect that there cannot be a difference in epistemic justification if there is not a difference in mental states36. As Conee & Feldman put it:

...if any two possible individuals are exactly alike mentally, then they are alike justificationally, e.g., the same beliefs are justified for them to the same extent (2001, p. 56).

34 An argument that Otto does base his belief on what propositionally justifies his belief, at T, would presumably have to presuppose something like the causal account of the basing relation. This faces well-known hurdles, such as Lehrer’s (1971) case of the gypsy lawyer. Moreover, it seems that if a doxastic account of the basing relation were correct, it’s implausible Otto could base his belief on what propositionally justifies his belief at T; in Leite’s (2004) parlance, Otto will not be able to justify his belief at T, and ipso facto, this undermines the basis of Otto’s belief.

35 Or, alternatively, unless one wants to go against orthodoxy and deny, as Kvanvig puts it, that ‘Doxastic justification is what you get when you believe something for which you have propositional justification, and you base your belief on that which propositionally justifies it.’

36 See Turri (2009) for a helpful discussion here.
Now, the natural reason for thinking the extended mind thesis stands at odds with this proposal is simple: by allowing for extended mental states—for instance, extended beliefs—the extended mind thesis will diffuse mental states as Clark (2010) puts it ‘firmly outside the head’ and in to the world in such a way that will not respect the mentalist’s supervenience thesis.

But why think this? The most natural reason to do so will be a background knee-jerk commitment to epistemic individualism\(^ {37} \): The thesis that defeasible epistemic justification supervenes on internal features of the biological agent. Indicatively epistemic individualism is actually a common presupposition of proponents of both internalism and externalism\(^ {38} \) alike in epistemology\(^ {39} \).

Obviously, with epistemic individualism in the background, we think of mental states of the sort at play in the mentalist’s supervenience thesis as essentially organismic and thereby non-extended\(^ {40} \) and then we reason that, in so far as a proponent of the extended mind thesis will have to accept that justification supervenes on properties of mental states that are extended, then the proponent of the extended mind thesis will be in strict tension with the mentalist internalist’s supervenience claim.

Of course, the proponent of the extended mind thesis rejects \textit{ex ante} the claim that mental states are essentially non-extended. This is important for the purpose of assessing the issue of whether were the extended mind thesis true, it would be incompatible with the mentalist internalist’s supervenience claim. Put another way: If we begin by assuming that mental states such as beliefs can be extended in the sense claimed by the proponent of the extended mind thesis, then \textit{must we reject} the claim that there is no difference in epistemic justification unless there is a difference in mental states?

In principle, no. Consider that there is logical space for carving out the following kind of position, which we can call epistemically anti-individualistic, or extended, mentalist-J internalism:

\[
\text{(Anti-individualistic) extended mentalist J-internalism: (i) if any two possible individuals are exactly alike mentally, then they are alike justificationally; (ii) the mental states in virtue of which two agents count as mentally alike can be both extended and non-extended.}
\]

\(^{37}\) For some outspoken opponents of this thesis, see Kallestrup & Pritchard (2013) and also Goldberg (2010; 2011).

\(^{38}\) As Pritchard & Kallestrup note, regarding Goldman: “According to [Goldman’s process reliabilism], a belief is justified just in case the (type of) cognitive process through which it was formed and sustained is reliable. Importantly, Goldman is explicit that the implicated processes are purely \textit{internal} – they are operations of the cognitive faculties residing inside the agent’s skin and skull” (Pritchard & Kallestrup, 87, our emphasis).

\(^{39}\) Sosa (2007, 29), a committed externalist, who defines knowledge in terms of getting to the truth because of intellectual virtue or competence, remarks that ‘... a disposition, one with its basis resident in the competent agent, one that would in appropriately normal conditions ensure (or make highly likely) the success of any relevant performance issued by it’. For some remarks where Goldman appears to commit himself to epistemic individualism, see Goldman (1979, 13) and Goldman (1986, 51). See also here Kallestrup \& Pritchard (2012, 100, fn. 11) for further discussion of Goldman’s commitment to epistemic individualism; see also (2013, pp. 86-87). Moreover, as Pritchard \& Kallestrup remark, robust virtue epistemology is normally characterised in such a way that it is clearly committed to epistemic individualism (\textit{Ibid}, p. 87).

\(^{40}\) For instance as ‘internal to consciousness’—or along such lines.
Of course, extended mentalist J-internalism is a thesis that the proponent of the extended mind thesis will take to be implied by the more basic mentalist-J internalist thesis, given that (ii) is implied by the extended mind thesis. Consider after all that it is precisely because the proponent of the extended mind thesis takes extended mentalist J-internalism to be implied by mentalist J-internalism that the proponent of the extended mind thesis and one who antecedently denies this thesis will disagree intractably about whether Otto and Otto* are in the same mental state at $T^4$. The upshot, then, is that premise (1) simply wouldn’t non-question-beggingly get off the ground.

This being said, not all sympathetic to mentalist J-internalism will be willing to sign on to extended mentalist J-internalism as an example of bona fide mentalist internalism. But this is because they will think that mentalist J-internalism entails a thesis we can call individualistic mentalist J-internalism:

Individualistic Mentalist J-internalism: (i) if any two possible individuals are exactly alike mentally, then they are alike justificationally; (ii) the mental states in virtue of which two agents count as alike mentally can only be non-extended mental states.

Individualistic mentalist J-internalism is ordinarily taken to be the thesis proffered by proponents of mentalist J-internalism. Notice though that mentalist J-internalism, as such, does not entail individualistic mentalist J-internalism in the absence of a further commitment to an individualist interpretation of mental states, as captured in claim (ii). Crucially, then, it is not in virtue of one’s being a mentalist J-internalist, as such, that one will not recognise extended mentalist J-internalism; it is because, rather, one is an epistemic individualist. This much vindicates the previous suggestion that extended mentalist-J internalism is a bona fide variety of mentalist J-internalism, which counts squarely against the incompatibility claim made in (iv).

An individualist mentalist J-internalist might grant this point but claim it to be a Pyrrhic victory on the grounds that extended mentalist J-internalism is so far removed from the traditional individualistic construal of the mentalist J-internalist thesis, that it does not count as a bona fide version of epistemic internalism. In a sense, we are happy to concede this point, but we think nothing important turns on it. In so far as epistemic anti-individualism is very different from epistemic individualism, the observation made here is surely true.

But, there’s no good reason to suppose that somehow the spirit of internalism lends itself only to an individualistic construal, simply because it is typically given such a construal. The guiding idea of epistemic internalism, after all, is upheld even by an extended mentalist J-internalist who parts ways with an epistemically anti-individualist externalist (such as a reliabilist) about epistemic justification. Just consider that an externalist of this sort will endorse the idea of extended mental states but, crucially, he will maintain that justification is a matter of the reliability of the processes issuing the extended mental states, which reliability may at least partly supervene on matters external to one’s mental life. In other words, he will reject the claim

---

41 This is precisely an argument for the incompatibility of mentalist J-internalism and the extended mind thesis (blatantly) could not be cogently established by starting with a premise where individuals like Otto and Otto* counts as being in the same mental states, in order to show (then) that Otto and Otto* diverge vis-à-vis their justification.
that justification is solely a matter of how things stand \textit{vis-à-vis} the agent’s mental states themselves (extended, or otherwise). This latter claim, however, is precisely what makes mentalism an internalist view even when it is stripped off the otherwise tacitly assumed individualism.

5 Conclusion

We therefore see that none of the claims (i)-(iv) is obviously true, let alone their conjunction, which constitutes the incompatibility thesis and which could count as at least \textit{prima facie} bad news for active externalism. Admittedly, this is not an entirely unsurprising conclusion to draw so it is fair to wonder why it may be easy to assume that at least some ways of combining active externalism with epistemic internalism would be obviously incorrect.

One possible explanation for this, as indicated by the closing lines of the previous section, concerns the (somewhat striking) shape that epistemic internalism would assume on active externalist construals. Consider, to begin with, the combination of accessibilist J-Internalism with the extended cognition hypothesis. When we move from denying the in-principle incompatibility of these positions to actively envisioning their union, we can conceive of the following kind of possible view: with the provision of an extended analogue to reflection in play, so long as one is able to access one’s (extended or non-extended) processes that currently justify one’s belief that $p$, one is justified in holding $p$ (no matter whether the relevant justificatory process is external to one’s organism). Call this \textit{epistemically anti-individualist, or extended access-process J-internalism}.

Similarly, the combination of accessibilist J-Internalism with the extended mind thesis, again with an extended analogue of reflection in hand, would allow that so long as one is able to access the (extended or non-extended) beliefs (regardless of whether they are dispositional or occurrent) that justify one’s belief that $p$, one is justified in holding $p$ (no matter whether the relevant justificatory mental states are external to one’s organism). Call this \textit{epistemically anti-individualistic or extended access-state J-internalism}.

Now, the envisaged combination of mentalist J-internalism with the extended cognition hypothesis turns out to be a kind of vacuous (or, perhaps, orthogonal) mix; even if we allow one’s cognitive processes to extend, this will have no notable effect on one’s justificatory status as, according to mentalism simpliciter, one is justified only with reference to one’s mental states. Call this ineffectual combination (for lack of a better term) \textit{epiphenomenal epistemically anti-individualistic, or extended, or J-mentalist internalism}.

As noted before, however, the combination of the extended mind thesis with mentalist J-internalism—\textit{viz.} \textit{extended J-mentalist internalism}—gives way to the interesting result that one can be justified in holding that $p$ on the basis of extended mental states; that is on the basis of states that may lie partly outside one’s organism.

No doubt, all the preceding forms of epistemic internalism are quite peculiar to say the least. Despite their lack of appeal, however, they are all worthy of the internalist label. Save epiphenomenal extended J-mentalist internalism, which is a view with no teeth, so to speak,
both extended access-process J-internalism and extended access-state J-internalism put forward the traditionally internalist requirement that one have, at least in principle, reflective access to one's justification for \( p \). And as indicated in the previous section, extended J-mentalist internalism too qualifies as epistemically internalist since it requires that one's justification be solely a matter of what is internal to one's mental life (even if one's mental life may extend beyond one's organism).

We therefore see that none of the above theses is objectionable on account of running against the spirit of epistemic internalism. True, they are all epistemically anti-individualist, but this is not the same as being externalist. To think otherwise would be to identify epistemic internalism with epistemic anti-individualism, a traditionally tacit assumption which now seems entirely unwarranted.

In conclusion, then, active externalism is far from being in principle incompatible with epistemic internalism per se and none of the two versions we have here considered significantly restricts one's options in epistemology, despite initial appearances to the contrary. In fact, it only seems to be incompatible with epistemic individualism but, given that active externalism can be defined as the denial of the sister thesis of cognitive individualism (i.e., the idea that one's cognitive economy is restricted within one's organism) this should hardly come as a surprise at all.\(^{42}\)

References


\(^{42}\) The authors would like to thank Emma C. Gordon, Jesper Kallestrup and Duncan Pritchard for helpful discussion. Thanks also to two anonymous referees at *Erkenntnis* who offered a number of very helpful suggestions. This article was written as part of the AHRC-funded 'Extended Knowledge' (#AH/J011908/1) research project that is hosted by the University of Edinburgh’s Eidyn research centre.


University Press.


