

Access Internalism and the Extended Mind

Declan Smithies

Forthcoming in *Extended Epistemology*, Oxford University Press

Edited by A. Carter, A. Clark, J. Kallestrup, O. Palermos, and D. Pritchard

Is internalism in epistemology compatible with externalism in philosophy of mind? The question is worth asking because the labels ‘internalism’ and ‘externalism’ have different meanings in epistemology and philosophy of mind. But the issue is further complicated by the fact that these labels can be used in several different ways even within each of these areas.

Internalism in epistemology can mean either *mentalism* or *accessibilism*. Mentalism is the thesis that which propositions one has justification to believe supervenes upon one’s mental states. Accessibilism is the thesis that one has privileged epistemic access to which propositions one has justification to believe. Neither mentalism nor accessibilism logically entails the other: these two claims can be combined as parts of a larger package, but either one can be consistently held without the other.¹

Externalism in philosophy of mind can mean either *content externalism* or *vehicle externalism*. Content externalism is the thesis that some mental states have their contents in virtue of relations to the external environment. Vehicle externalism is the thesis that some mental states – that is, vehicles of mental content – are partially constituted by features of the external environment. Again, neither content externalism nor vehicle externalism logically entails the other.²

¹ The terms ‘mentalism’ and ‘accessibilism’ are from Conee and Feldman 2001. See Conee and Feldman 2001 for mentalism without accessibilism, Gibbons 2006 for accessibilism without mentalism, and Smithies 2014 for the combination of mentalism and accessibilism.

² The term ‘vehicle externalism’ is from Hurley 1998. Burge 1986: 13-18 combines content externalism with vehicle internalism: beliefs are constituted by neural states but have their contents in virtue of relations to the external environment. Clark and Chalmers 1998: 14, n. 4 combine vehicle externalism with content

The question whether epistemic internalism is compatible with content externalism has been widely discussed. In contrast, the question whether epistemic internalism is compatible with vehicle externalism has been largely neglected until recently.³ This is perhaps not surprising given that content externalism emerged in the 1970s, whereas vehicle externalism did not emerge until the 1990s. Even so, the time is ripe for addressing the second question. First, though, we need to clarify what kind of epistemic internalism we are talking about. It seems quite clear that mentalism is compatible with vehicle externalism so long as the mental states on which justification supervenes can be partially constituted by the environment. A more interesting question is whether accessibilism is compatible with vehicle externalism. That is the target question for this chapter.

The main goal of this chapter is to argue that accessibilism is incompatible with vehicle externalism. As we shall see, however, there are strong arguments for both of these positions. On the one hand, there is a compelling argument for vehicle externalism: the parity argument from Clark and Chalmers 1998. On the other hand, there is a compelling argument for accessibilism: the Moorean argument from Smithies 2012. If accessibilism is incompatible with vehicle externalism, then both arguments cannot be sound. I resolve the tension by arguing that the Moorean argument succeeds, while the parity argument fails, and hence that vehicle externalism should be rejected on broadly epistemological grounds.

Here is the plan for the paper. Section 1 presents the parity argument for vehicle externalism. Section 2 presents the Moorean argument for accessibilism. Section 3 presents a new argument that accessibilism is incompatible with vehicle externalism. Sections 4-6 defend the premises of this argument against objections. Section 7 explains why the parity argument for vehicle externalism fails and draws some more general lessons about functionalism and the mind. Section 8 concludes.

internalism: beliefs have “notional” contents in virtue of features that are internal to a cognitive system that is extended into the external world.

³ See Carter, Kallestrup, Palermos and Pritchard 2014 for an overview of the current state of the literature on both questions. See also Carter and Palermos 2015 for a recent discussion of the second question.

1. Extended Cognition

I will focus my discussion on the version of vehicle externalism proposed by Andy Clark and David Chalmers in their groundbreaking paper, “The Extended Mind”. They argue that beliefs can be extended in the sense that they are constituted in part by aspects of the external environment that play the right kind of functional role in a subject’s cognitive processes. I’ll call this the *extended cognition thesis*. Here is an official statement of the thesis from the original paper:

The extended cognition thesis: “beliefs can be constituted partly by features of the environment, when those features play the right sort of role in driving cognitive processes.” (Clark and Chalmers 1998: 12)

I’ll begin by clarifying the thesis before presenting the central argument for it.

First, this is a thesis about cognition as distinct from its computational underpinnings. As I use the term, cognition is distinguished from mere computation by the fact that it is subject to epistemic evaluation. By this criterion, cognition includes doxastic states of belief and doxastic processes of belief revision, but it excludes the subdoxastic states and processes that figure in computational explanations in cognitive science.

Second, this is a thesis about cognition as distinct from consciousness. Clark and Chalmers concede that it is “far from plausible that consciousness extends outside the head” (1998: 10). But cognition may be extended even if consciousness is not. Moreover, this is compatible with fairly strong connections between consciousness and cognition. Following Clark and Chalmers, I will assume that beliefs are standing dispositional states, rather than occurrent events in the stream of consciousness. But this is consistent with (although it does not require) the view that beliefs are dispositions to cause occurrent events of consciously thinking or judging the contents that one believes.

Third, Clark and Chalmers argue for the possibility of extended cognition, but they do not argue – at least, not in the original paper – that this possibility is actual. Their argument is based on consideration of counterfactual examples, rather than

actual ones – in particular, the example of Inga and Otto discussed below. But if their argument is sound, then it has the potential to be extended to actual examples. Clark (2011) argues for this extension in some detail.

Finally, the extended cognition thesis is more controversial than the thesis that beliefs can have an extracranial location and a non-neural basis.⁴ Given the commitment to multiple realizability, many functionalists would accept that beliefs can be realized by implanting silicon chips in the brain or by wiring them up to the brain remotely. Even so, it is much more controversial that beliefs can have an extracranial location and a non-neural basis just by playing the kind of functional role that Otto's notebook plays in Clark and Chalmers' example. One can deny that Otto's notebook realizes extended beliefs without thereby ruling out the possibility of more remote science-fiction scenarios in which beliefs have an extracranial location and a non-neural basis.

Let us turn now to the key example of Inga and Otto. Clark and Chalmers describe the case as follows:

Inga hears from a friend that there is an exhibition at the Museum of Modern Art, and decides to go see it. She thinks for a moment and recalls that the museum is on 53rd Street, so she walks to 53rd Street and goes into the museum. It seems clear that Inga believes that the museum is on 53rd Street, and that she believed this even before she consulted her memory. It was not previously an *occurrent* belief, but then neither are most of our beliefs. The belief was sitting somewhere in memory, waiting to be accessed.

Now consider Otto. Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he

⁴ Compare Farkas (2012) for two versions of the extended mind thesis: (i) "the physical basis of mental events can extend beyond the boundaries of our physical body" (2012: 437); and (ii) "the typical role of standing states can be extended to include states that produce conscious manifestations in a somewhat different way than normal beliefs and desires do" (2012: 441).

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. Today, Otto hears about the exhibition at the Museum of Modern Art, and decides to go see it. He consults the notebook, which says that the museum is on 53rd Street, so he walks to 53rd Street and goes into the museum. (1998: 12-13)

Why do Inga and Otto walk to 53rd St? According to Clark and Chalmers, the explanation is the same in both cases: each of them wants to go to MoMA and believes that MoMA is on 53rd St. The key point is that Otto's notebook plays the same role for him that Inga's memory plays for her – namely, combining with desire to produce action. This is often thought to be a central role – and perhaps even *the* central role – in terms of which belief is defined.⁵

One might object that Otto has no beliefs about the location of MoMA until he consults his notebook. If so, then the explanation of his behavior is more complicated. Otto walks to 53rd St because he wants to go to MoMA, and he believes that MoMA is located wherever it says in the notebook, but he does not believe that MoMA is located at 53rd St until he opens his notebook and looks at it.

Clark and Chalmers reply that we should avoid explaining Otto's behavior in this way because it adds pointless complexity. First, they argue that we shouldn't complicate the explanation in Otto's case unless we do so in Inga's case too. If we say that Otto has no belief that MoMA is on 53rd St until he consults his notebook, then we should say that Inga has no such belief until she consults her memory. Second, they argue that, other things being equal, simpler explanations are better than more complicated ones: "In an explanation, simplicity is power" (1998: 13).

If Otto's notebook plays the same role for him that Inga's memory plays for her, then we can argue that Otto's notebook realizes belief if Inga's memory does. Here is a reconstruction of *the argument from parity*:

⁵ Compare Stalnaker: "To believe that *P* is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in a world in which *P* (together with one's other beliefs) were true" (1984: 15).

- (1) Anything that plays the same role as a belief is itself a belief.
- (2) The information stored in Inga's memory plays the right kind of role to count as belief.
- (3) The information stored in Otto's notebook plays the same role as the information stored in Inga's memory.
- (4) Therefore, the information stored in Otto's notebook plays the right kind of role to count as belief.

The argument has three premises. I'll briefly consider each of them in turn.

The first premise is a version of what has become known as *the parity principle*. In the first instance, Clark and Chalmers define the parity principle for cognitive processes:

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. (1998: 8).

To paraphrase, if some process plays the same functional role as a cognitive process, then it is a cognitive process. The same principle applies to cognitive states: if an informational state plays the same functional role as a belief, then it is a belief. This principle is motivated by a broadly functionalist conception of belief, according to which a belief is an informational state that plays the right kind of functional role. As Clark and Chalmers put the point, "What makes some information count as a belief is the role it plays" (1998: 14).

One might object that functionalism is false. But the parity principle requires only a weak form of functionalism. There is no commitment to functionalism about consciousness as distinct from cognition. Moreover, there is no commitment to any specific account of the functional role of cognition. For instance, the functional role of belief might include, and might even be exhausted by, its disposition to cause

conscious experiences of thinking and judging that are immune from any further functional analysis.⁶

One might object that even this weak form of functionalism is false. Thus, Mark Sprevak (2009) argues that the extended cognition thesis reduces to absurdity the functionalist assumptions that motivate the parity principle. But it is hard to see how this weak form of functionalism could be false. What more is needed for belief besides having an informational state that plays the right kind of functional role? One might insist that belief requires neural realization or intracranial location, but this conflicts with the plausible and widespread assumption that cognition can be multiply realized.

The rationale for the second premise is that if anyone believes anything, then Inga does, since a paradigm case of believing that *p* is having the information that *p* stored in memory. One might object that there is no such mental state as belief construed as a standing, dispositional state, rather than an occurrent event in the stream of consciousness. According to Brie Gertler, for instance, “the mind is made up entirely of *occurrent* states and *conscious* processes.” (2007: 203). There is no room for standing, dispositional belief within this conception of the mental. On this view, there is parity between Otto and Inga because neither has a mental state of believing that MoMA is on 53rd St.

On the face of it, this view is extremely revisionary. The existence of dispositional belief is supported not only by ordinary language, but also by explanatory and epistemic considerations. Suppose that I open the fridge while making my morning coffee. What explains my behavior? Let’s stipulate that I’m so preoccupied with my plans for the day that I don’t consciously think about what I’m doing at all. Even so, it would be perfectly natural to say that I open the fridge because I want milk in my coffee and I believe there is milk in the fridge. It’s not only part of common sense that such belief ascriptions are often true. It’s hard to explain why I act as I do, and why my action is rational or justified, without assuming that I am acting on standing beliefs that are not occurrent in conscious experience.

⁶ See Schwitzgebel 2002 and Smithies 2014, 2016 for versions of this view.

The rationale for the third premise is that Otto's notebook plays the same role in explaining his behavior that Inga's memory plays in explaining hers. One might object that there are differences between Inga and Otto with regard to the reliability or stability of their access to stored information. For instance, Clark and Chalmers note that Otto cannot access his notebook in the shower or in the dark. But they also note that Inga cannot access her memory when she is asleep or intoxicated. Arguably, then, parity remains.

Clark and Chalmers stipulate that Otto's access to his notebook is no less stable or reliable than Inga's access to her memory. Moreover, they stipulate that it is only under these conditions that Otto's notebook constitutes a belief. If his access to the notebook is too unstable or too unreliable, then it does not constitute a belief at all. According to Clark and Chalmers, the following conditions must be satisfied:

- (1) First, the notebook is a constant in Otto's life – in cases where the information in the notebook would be relevant, he will rarely take action without consulting it.
- (2) Second, the information in the notebook is directly available without difficulty.
- (3) Third, upon retrieving the information from the notebook he automatically endorses it.
- (4) Fourth, the information in the notebook has been consciously endorsed at some point in the past, and indeed is there as a consequence of this endorsement. (1998: 17)

The fourth condition is controversial because it contains a historical component. Arguably, I can believe things that I have never endorsed in the past so long as I am now disposed to endorse them when the question arises. But the historical condition can be replaced with the causal condition that information is typically stored in the notebook whenever it is consciously endorsed.

Even granting these functional similarities between Inga and Otto, one might insist that there are some functional differences that remain.⁷ But the challenge is to make the case that these functional differences are deep and important enough to make for a cognitive difference between Inga and Otto. After all, Clark and Chalmers can allow for superficial differences between Inga and Otto:

The differences between Otto's case and Inga's are striking, but they are superficial. By using the "belief" notion in a wider way, it picks out something more akin to a natural kind. The notion becomes deeper and more unified, and is more useful in explanation. (1998: 14)

As Sprevak (2009) notes, if cognition is multiply realizable, then we should expect fine-grained differences in functional role corresponding to differences in physical realization. So the challenge is to make the case that the functional differences between Inga and Otto are coarse-grained enough to make for a cognitive difference, rather than merely a difference in how cognition is realized. One of the aims of this paper is to develop a new response to this challenge.

2. Accessibilism

Accessibilism is the thesis that one always has privileged access to the facts about which propositions one has justification to believe. There are various different ways of making this privileged access thesis more precise, but in Smithies 2012 I defined it as a biconditional version of the JJ principle:

Accessibilism: one has justification to believe that p if and only if one has justification to believe that one has justification to believe that p ($Jp \leftrightarrow JJp$).

According to accessibilism, the facts about which propositions one has justification

⁷ For specific proposals, see Adams and Aizawa 2001, Rupert 2004, Gertler 2007, Weiskopf 2008, and Pearlberg and Schroeder 2015.

to believe are *self-intimating* in the sense that one has higher-level justification to believe that they obtain whenever they do obtain. Moreover, one's higher-level justification is *infallible* in the sense that these justification facts obtain whenever one has higher-level justification to believe that they obtain. In that sense, one has privileged epistemic access to the facts about justification.

As I have defined it, accessibilism is a thesis about *propositional justification*, rather than *doxastic justification*: it is a thesis about which propositions one has justification to believe, rather than which justified beliefs one has. Doxastic justification requires propositional justification, but not vice versa: one can have justification to believe a proposition that one does not believe in a way that is justified or that one does not believe at all. Arguably, one can have justification to believe a proposition even if one does not have the psychological capacity to believe it in a way that is justified. For instance, one might be too tired, too distracted, or just too dense to believe whatever is justified by the evidence.

This is crucial for avoiding some stock objections to accessibilism. Accessibilism does not entail that having justified beliefs requires having justified higher-order beliefs about those beliefs. That would rule out the possibility of justified belief in children and animals with a capacity for cognition but not metacognition. Worse still, it rules out the possibility of justified belief in any finite agent, since having justified beliefs would require having an infinite hierarchy of justified higher-level beliefs of infinitely increasing complexity. Since accessibilism is a thesis about propositional justification, rather than doxastic justification, it does not have these implausible consequences.

In Smithies 2012, I argued for accessibilism by appealing to the absurdity of believing Moorean conjunctions of the following forms:

- (1) p but I don't have justification to believe that p
- (2) p but it's an open question whether I have justification to believe that p
- (3) I have justification to believe that p , but it's not the case that p
- (4) I have justification to believe that p , but it's an open question whether p

The argument takes the form of a reductio. If accessibilism is false, then I can have justification to believe Moorean conjunctions of these forms. Intuitively, however, I cannot have justification to believe these Moorean conjunctions because believing them would be absurd. Therefore, accessibilism is true.

Consider the left-to-right direction first. If it is false, then there is a possible case in which I have justification to believe that p , while lacking justification to believe that I have justification to believe that p . But now we can ask: which attitude do I have justification to adopt towards the proposition that I have justification to believe that p ? If I don't have justification to believe it, then I have justification either to disbelieve it or to withhold belief. So I have justification to believe that p , while also having justification to disbelieve or to withhold belief in the proposition that I have justification to believe that p . Assuming that justification is closed under conjunction, it follows that I have justification to believe (1) or (2).

Now consider the right-to-left direction. If it is false, then there is a possible case in which I have justification to believe that I have justification to believe that p , but I lack justification to believe that p . But now we can ask: which attitude do I have justification to adopt towards the proposition that p ? If I don't have justification to believe it, then I have justification either to disbelieve it or to withhold belief. So I have justification to believe that I have justification to believe that p , while also having justification to disbelieve or to withhold belief in the proposition that p . Assuming that justification is closed under conjunction, it follows that I have justification to believe (3) or (4).

There are various options for blocking the argument, but none of them has much plausibility. One option is to bite the bullet and insist that I can have justified belief in Moorean conjunctions. But I cannot *know* the Moorean conjunction that p and I don't have justification to believe that p , for if I know the first conjunct, then I have justification to believe it, and so the second conjunct is false. Moreover, I can know that I cannot know the Moorean conjunction. And it's not plausible that I can have justification to believe a proposition that I can know I cannot know.

Another option is to say that I can have justification to believe Moorean conjunctions, but I cannot have justified belief because my justification is *finkish* in

the sense that it is destroyed in the act of trying to use it. More specifically, one might say that my justification to believe that p is destroyed whenever I disbelieve or withhold belief that I have justification to believe that p . But it's not plausible that my justification to believe that p is destroyed when these higher-order attitudes are unjustified. It's only when I have justification for these higher-order attitudes that my first-order justification is destroyed. But that's just what accessibilism implies: I have justification to believe a proposition only if I have justification to believe that I have justification to believe it. So it's hard to sustain the view that I have finkish justification for believing these Moorean conjunctions.

A third option is to reject the principle that justification is closed under conjunction. But even if justification does not *always* transmit across a conjunction, as the preface paradox suggests, it does not follow that justification *never* transmits across a conjunction. So the objection does not rule out the possibility of having justification to believe Moorean conjunctions. In any case, one cannot avoid the charge of irrationality just by failing to conjoin one's beliefs in the conjuncts of a Moorean conjunction. So, the argument can be run using only the weaker premise that one cannot have justification to believe the conjuncts of a Moorean conjunction.

I conclude that the argument for accessibilism is sound. But is accessibilism a form of internalism? It is often assumed so. Arguably, however, some externalist theories of justification are consistent with accessibilism. To illustrate the point, consider the (admittedly implausible) theory that one has justification to believe all and only truths.⁸ If it is true that p , then one has justification to believe that p , and if that is true, then one has justification to believe that one has justification to believe that p , and so on. So this theory of justification implies accessibilism. Arguably, however, it deserves to be categorized as an externalist theory of justification. If so, then how should we understand the distinction between internalist and externalist versions of accessibilism?

An internalist form of accessibilism – call it *access internalism* – combines a privileged access thesis with a peculiar access thesis. Access internalism is the thesis

⁸ I owe this example to Jonathan Jenkins Ichikawa.

that our epistemic access to facts about justification is not only privileged, but also peculiar: it is *better* than other modes of epistemic access, but also *different* from other modes of epistemic access. More precisely, our epistemic access to facts about justification has its source in reflection alone – that is, it has its source in a combination of introspection and a priori reasoning. So access internalism can be defined as follows:

Access Internalism: one has justification to believe that *p* if and only if one has justification to believe *through reflection alone* that one has justification to believe that *p*.

An externalist form of accessibilism – call it *access externalism* – accepts the privileged access thesis but rejects the peculiar access thesis. I'll return to this distinction between access internalism and access externalism in due course.

3. An Argument for Incompatibilism

In sections 1 and 2, we saw that there are plausible arguments for the extended cognition thesis and for accessibilism. However, I will now argue that these two theses are incompatible and so cannot both be true: if accessibilism is true, then the extended cognition thesis is false, and vice versa. Here is the basic argument for incompatibilism:

- (1) If accessibilism is true, then anything that plays a justifying role is accessible through introspection.
- (2) All beliefs play a justifying role.
- (3) If the extended cognition thesis is true, then not all beliefs are accessible through introspection.
- (4) Therefore, if accessibilism is true, then the extended cognition thesis is false.

In sections 4-6, I'll defend each premise of the argument against objections, but I'll begin by briefly explaining the motivation for each premise.

The rationale for the first premise is that accessibilism is best explained by access internalism rather than access externalism. Access internalism explains our privileged access to facts about justification in terms of our privileged access to the mental facts that determine them. The justification facts are determined by mental facts that are accessible through introspection. Moreover, the justification facts are determined by mental facts in accordance with epistemic principles that are accessible through a priori reasoning. This guarantees that the justification facts are accessible through reflection alone – that is, through a combination of introspection and a priori reasoning.

The first premise is rejected by proponents of *access externalism* who endorse accessibilism while rejecting access internalism (Gibbons 2006). On this view, accessibilism can be true even if not everything that plays a justifying role is accessible through introspection. In particular, the information stored in Otto's notebook can play a justifying role even if it is not accessible through introspection. I'll argue in section 4 that we need to endorse access internalism rather than access externalism in order to explain the truth of accessibilism.

The rationale for the second premise is that the role we associate with the ordinary concept of belief includes its epistemic role in justifying other beliefs. All beliefs have the potential to justify other beliefs. Not all beliefs fulfill this potential, since only justified beliefs justify other beliefs. But all beliefs have the potential to justify other beliefs when they themselves are justified. Indeed, this is part of what sets beliefs apart from subdoxastic states, which have no such potential to play an epistemic role in justifying and being justified.

The second premise is rejected by proponents of *strong phenomenal mentalism* who claim that only conscious experience can play a role in justifying belief (Conee and Feldman 2008). On this view, dispositional beliefs cannot play a justifying role. So neither Inga's memory nor Otto's notebook stores information in a way that plays a justifying role. I'll argue in section 5 that this proposal fails to explain some key features of the epistemology of ordinary reasoning.

The rationale for the third premise is that Otto's access to the information stored in his notebook relies upon perception, rather than introspection. Inga can

know that she believes that MoMA is on 53rd St just by consulting her memory, whereas Otto has to open his notebook and read it. Therefore, if Otto has extended beliefs that are partially constituted by the information stored in his notebook, then those extended beliefs are not introspectively accessible.

The third premise is rejected by proponents of *extended introspection* who claim that Otto's perception of his notebook constitutes an extended form of introspection (Clark and Chalmers 1998). I'll argue in section 6 that this proposal obscures an important epistemic asymmetry between perception and introspection.

4. Access Externalism

The first compatibilist option is to reject access internalism in favor of access externalism. The term 'access externalism' was coined by John Gibbons (2006). According to Gibbons' version of access externalism, justification supervenes upon what you're in a position to know, but it doesn't supervene upon what you're in a position to know *through reflection alone*.

If Gibbons is right, then the information stored in Otto's notebook need not be introspectively accessible in order to play a justifying role. What matters is that Otto is in a position to know the contents of his notebook by opening and reading it. If so, then the information stored in his notebook can play a justifying role even if it is not accessible to him through introspection.

Gibbons argues against access internalism by proposing pairs of subjects who are alike in what they can know by introspection, but who differ in what they have justification to believe owing to differences in what they can easily come to know by perception. Here is his main example:

The good case: I have justification to believe that I'll have eggs for breakfast because I seem to remember seeing eggs in the fridge last night.

The bad case: I lack justification to believe that I'll have eggs for breakfast, although I seem to remember seeing eggs in the fridge last night, because I should have seen the sign on the door that reads, "We're out of eggs!"

Gibbons argues that in the bad case (1) I should have known we're out of eggs because I should have seen the note on the door; and (2) I shouldn't believe what I should know to be false; but (3) I don't have justification to believe what I shouldn't believe; so (4) I don't have justification to believe that I'll have eggs for breakfast.

In my view, the argument fails because it trades on an equivocation between different senses of 'ought' or 'should'.⁹ There is a subjective sense in which I should believe whatever is supported by the evidence that is currently in my possession. In the subjective sense, I should believe that I'll have eggs for breakfast. But there is also an objective sense in which I should believe whatever is supported by the evidence that I could easily come to possess by using better methods of gathering evidence. In the objective sense, I shouldn't believe that I'll have eggs for breakfast because I should have known better.

Richard Feldman (2004) draws a related distinction between two senses of epistemic rationality:

One sort of epistemic appraisal concerns whether believing a particular proposition is rational for a person at a time given exactly the situation the person happens to be in at the time. We may say that this is an assessment of the *current-state epistemic rationality* of believing the proposition.... A second possible epistemic evaluation of a belief has to do with the methods that led to it. We may call this *methodological epistemic rationality*. (2004: 233)

As Feldman notes, a belief can be rational in the current-state sense without being rational in the methodological sense: a belief can be supported by someone's current evidence even if they possess this evidence as a result of using bad methods. In the bad case, my belief that I'll have eggs for breakfast is justified by the evidence that is currently in my possession, although my possession of this evidence results from insufficiently careful methods of gathering evidence.

⁹ The difference is best modeled in Kratzer's (1981) semantics by means of a distinction between objective and subjective ways of ranking options.

Access internalism is best construed as a theory of epistemic rationality or justification in the current state sense, rather than the methodological sense. Gibbons' argument against access internalism fails for this reason. But we still need to ask whether Gibbons is right to claim that we can explain accessibilism without endorsing access internalism. If so, then much of the motivation for access internalism is undermined.

In section 2, I defined accessibilism as a biconditional version of the JJ thesis. Gibbons (2006: 30) explicitly accepts the right-to-left direction of the biconditional. He argues for this conditional version of the JJ thesis, as I have done, by appealing to the impermissibility of believing Moorean conjunctions:

From your point of view, the question of whether you are justified in believing that p is not independent of the question of whether p is true. If the questions were independent, any pair of answers would be acceptable in some situation or another. But you should never believe a proposition of the form (p but I am not justified in believing that p). Nor should you believe anything of the form (p is false, but I am justified in believing it). (2006: 32-3)

As we saw in section 2, however, this Moorean argument can be extended to both directions of the biconditional. Gibbons does not explicitly accept the left-to-right direction of the biconditional, but he comes very close:

In fact, it is tempting to suppose that first-order justification itself is directly relevant to the justification of the relevant second-order belief if you form one. In the ordinary case, your second-order belief is justified because your first-order belief is justified. (2006: 33)

As far as I can tell, his only qualm about the left-to-right direction stems from not distinguishing sharply enough between propositional and doxastic justification. If the topic is propositional justification, then first-order justification is sufficient for second-order justification whether or not you form a second-order belief. So I'll

continue to assume that a satisfying form of access externalism needs to explain the biconditional version of the JJ thesis.

In the following passage, Gibbons proposes an externalist explanation of our privileged access to facts about justification:

There is nothing mysterious about combining externalism about justification with the sort of privileged access that I have identified. One way to read (JJ) is to think of it as the claim that first-order justification is a condition for second-order justification. If any external facts are relevant to the determination of first-order justification, those very same facts are relevant to the determination of your second-order justification. (2006: 35)

The proposal is that the JJ thesis is true because anything that is relevant to the determination of first-order justification is also relevant to the determination of second-order justification too. If this is true, then it is sufficient to explain both directions of the biconditional. But does Gibbons explain why this is true?

On Gibbons' theory, what you have justification to believe depends upon what you're in a position to know. In the bad case, you have justification to believe that you're out of eggs because you're in a position to know that you're out of eggs. Does it follow that when you're in a position to know that you're out of eggs, you also have second-order justification to believe that you have first-order justification to believe that you're out of eggs? Not unless we're assuming that you're always in a position to know what you have first-order justification to believe. But that would be to assume that you have a kind of privileged access to facts about justification, which is just what needs to be explained. So I conclude that Gibbons' access externalism fails to explain our privileged access to facts about justification.

I don't claim that access externalism is an inconsistent view. But I do claim that access internalism explains our privileged access to facts about justification, whereas it is not clear how access externalism can explain this. I therefore conclude that we should reject access externalism in favor of access internalism.

5. Strong Phenomenal Mentalism

The second compatibilist option is to deny the premise that all beliefs can play a justifying role on the grounds that only conscious experiences can play such a role. On this view, neither the information stored in Inga's memory nor the information stored in Otto's notebook plays a justifying role.

Suppose that access internalism is true and hence that anything that plays a justifying role is accessible through introspection. And suppose that the information stored in Otto's notebook is not accessible through introspection and so cannot play a justifying role. This is no objection to the extended mind thesis unless we accept that all beliefs can play a justifying role.

Earl Conee and Richard Feldman (2008) accept a version of *evidentialism* on which one's evidence determines which propositions one has justification to believe. They also endorse the thesis that one's evidence is restricted to one's current conscious experience. On this view, only conscious experience can play a role in determining which propositions one has justification to believe.

Feldman (2004) motivates this view by drawing a comparison between two cases. In each case, my friend Jones assures me that I can complete the hike to Precarious Peak, and I believe him. In the first case, I have information stored in memory that Jones tends to overestimate my abilities. And in the second case, I have information stored in my pocket guidebook that the hike is suitable only for expert climbers. So in both cases I have information that counts against Jones' assurance, but in one case it is stored in memory and in the other case it is stored in a pocket guidebook. Moreover, the same question arises in each case: namely, am I justified in my belief that I can complete the hike to Precarious Peak?

According to Feldman, the answer is the same in each case: my belief is justified in the *current-state* sense, but not in the *methodological* sense. On this view, my belief is supported by my current evidence, but my current evidence results from using bad methods for acquiring evidence. If I had used better methods – say, by consulting my memory or by reading my guidebook – then I would have had evidence that defeats my justification to believe that I can complete the hike. As things are, however, I don't have any such defeating evidence.

On this view, the information stored in my memory is no more part of my evidence than the information stored in my pocket guidebook. This has clear implications for the case of Inga and Otto: neither the information stored in Inga's memory nor the information stored in Otto's notebook is suited to play an epistemic role in justifying belief.

Feldman considers three objections to this proposal: the problem of stored beliefs, the problem of forgotten evidence, and the problem of background beliefs. I'll argue that his responses to these objections fail.

First, *the problem of stored beliefs*. You have beliefs stored in memory that are justified, and often constitute knowledge, but which are not justified by anything in your current stream of consciousness. Here is an example from Feldman:

For example, while listening to a philosophy lecture you still know, and are justified in believing, that Washington D. C. is the capital of the United States. But (presumably) that wasn't supported by what you were thinking of during the lecture. (2004: 236)

In response, Feldman draws a distinction between occurrent and dispositional senses of epistemic terms, such as knowledge, justification, or rationality. The claim is that my dispositional beliefs constitute dispositional knowledge because their occurrent manifestations would constitute occurrent knowledge:

A person knows a thing dispositionally provided the person would know it occurrently if he thought of it. Since the thought that Washington is the capital would, presumably, be accompanied by an awareness of justifying evidence, this fact can be known dispositionally by most of us. Hence, the intuition that we know simple facts even when we are not thinking of them can be accommodated by the minimalist view of evidence possessed: they are known dispositionally but not occurrently. (2004: 236)

This reply commits a version of the conditional fallacy, since there are propositions that I don't know now, although I would know them if I were to think of them. Consider, for example, the proposition that I'm now thinking about rhubarb. Feldman acknowledges this problem, but he attempts to circumvent it by opting for a mild form of skepticism about dispositional knowledge:

In the most fundamental sense, one does not know things such as that Washington is the capital when one is not thinking of them. (2004: 237)

On this view, we don't really know anything that we're not currently thinking about, although we're disposed to have such knowledge whenever we need it. One consequence of this proposal is that we cannot explain our occurrent knowledge in terms of our dispositional knowledge. As we'll see, this makes trouble for Feldman's responses to other objections.

Second, *the problem of forgotten evidence*. I can know that Washington is the capital even if I have forgotten the evidence on the basis of which I originally acquired this knowledge. In that case, there may be nothing stored in memory that would justify the belief when it is brought into consciousness.

Feldman's response is that my justification to believe that Washington is the capital has its source in "my current conviction or feeling of certainty" (2004: 238). But what justifies this feeling? Presumably, it stands in need of justification, since I can feel too strongly convinced that something is true. And a feeling of conviction cannot justify a belief unless it is justified. But it is not clear what justifies my feeling of conviction that Washington is the capital without invoking my dispositions to feel certainty or conviction about other things. This threatens to undermine Feldman's proposal because he is committed to denying that occurrent knowledge is explained in terms of dispositional knowledge.

Third, *the problem of background beliefs*. If my justifying evidence is exhausted by my current experience, then how do we account for the epistemic role of background beliefs? In Feldman's example, an expert and a novice see a scarlet tanager and it looks the same way to each of them. The expert comes to believe that

it's a scarlet tanager because she has the background knowledge that scarlet tanagers look that way. Meanwhile, the novice leaps to the conclusion that it's a scarlet tanager on the basis of her knowledge that scarlet tanagers have once been spotted here before. In this case, the expert and the novice have just the same experiences, but only the expert knows that it's a scarlet tanager. Moreover, it's hard to explain this without invoking background knowledge.

As Feldman acknowledges, it would be implausible to maintain that the expert cannot know that it's a scarlet tanager unless she brings her background knowledge into consciousness. There is no need for the expert to make the inference consciously once it has become automatic. Instead, Feldman appeals to the expert's feelings of certainty that it's a scarlet tanager. The problem is that the novice can have these feelings too, although they're not justified in his case. So it's difficult to explain the difference in epistemic standing between the expert and the novice without appealing to the epistemic role of background beliefs.¹⁰

The upshot is that we cannot plausibly restrict one's evidence to one's current conscious experience. One's dispositional beliefs also play a justifying role. But given access internalism, they can play this role only if they are accessible through introspection. The problem for the extended cognition thesis is that the information stored in Otto's notebook does not seem to satisfy this condition.

6. Extended Introspection

The third compatibilist option is to argue that the information stored in Otto's notebook is accessible to him through introspection. Clark and Chalmers take this line in response to an objection to the parity argument. The objection is that there is a disparity between Inga and Otto because Inga's access to her memory relies upon introspection, whereas Otto's access to his notebook relies upon perception. Clark and Chalmers reply that this objection begs the question:

¹⁰ Feldman (2004: 239) considers a reply that concedes this point. On this view, one's evidence includes not only conscious experience, but also background beliefs that are currently operative in causing one's current experience.

Otto's internal processes and his notebook constitute a single cognitive system. From the standpoint of this system, the flow of information between notebook and brain is not perceptual at all; it does not involve the impact of something outside the system. It is more akin to information flow within the brain. (1998: 16)

Presumably, Clark and Chalmers do not mean to deny that Otto perceives his notebook. That would be hard to swallow. A more charitable interpretation is that Otto's perception of the notebook constitutes an extended form of introspection, since it enables him to know the contents of his extended beliefs. The objection begs the question by assuming that if Otto's access to the notebook involves perception, then it cannot also constitute an extended process of introspection.¹¹

In some ways, this reply is too quick. Not all knowledge of one's beliefs counts as introspective knowledge – for example, knowledge that is acquired through the expert testimony of a therapist. So even if we assume for the sake of argument that Otto believes the contents of his notebook, it doesn't follow that his perception of the notebook constitutes a form of introspection. Still, it remains an open question whether Otto's perception of the notebook constitutes an extended form of introspection of his beliefs. We need an argument to settle the question.

Brie Gertler argues for a negative answer. She begins by noting that introspection is an exclusively *first-personal* way of acquiring knowledge: "it reveals only the introspector's own states, and not the states of others" (2007: 194). Next, she argues that Otto's way of acquiring knowledge of his beliefs through perception of his notebook does not satisfy this condition. After all, Otto's friend can figure out what Otto believes by reading his notebook in much the same way that Otto can. Thus, she writes:

When Otto tries to figure out what he believes on a particular topic, he consults the notebook. For instance, suppose that he wonders what he

¹¹ See also Carter and Palermos 2015: 760 for a similar point.

believes about the location of the MoMA. He will look in the notebook and conclude: I believe that the MoMA is on 53rd St. But of course someone other than Otto can determine Otto's beliefs in precisely the same way: by consulting the notebook, a friend can determine that Otto believes that the MoMA is on 53rd St. (2007: 194)

On the basis of this argument, Gertler concludes that Otto cannot acquire introspective knowledge of his beliefs by perceiving the notebook.

One might question Gertler's assumption we have some exclusively first-personal way of knowing about our beliefs. For instance, Clark argues that it is a purely contingent fact that other people cannot know what Inga believes by accessing her stored memories:

But why suppose that uniqueness of access is anything more than a contingent fact about standard biological recall? If, in the future, science devised a way for you to occasionally tap into my stored memories, would that make them any less *mine*, or part of my cognitive apparatus? (2010: 57)

As far as I can see, however, this objection misses the mark. Gertler does not rule out the possibility that other people can know what I believe by accessing my memories. What she rules out is the possibility that they can know what I believe by means of introspection. The futuristic possibility that Clark describes is one in which other people know what I believe by a novel form of perception, such as telepathy, rather than by introspection.

A more promising objection targets Gertler's assumption that Otto's friend can know what Otto believes in the same way that Otto can. If we're assuming that Otto's notebook plays the same role as Inga's memory, then Otto can know what he believes non-inferentially on the basis of perception of the notebook. Otto's friend, in contrast, cannot know that Otto believes what is written in his notebook without making an inference that relies on the background belief that the notebook plays the

right kind of role in Otto's psychology. Arguably, then, Otto's way of knowing what he believes is exclusively first-personal.

I propose a different argument for the conclusion that Otto cannot acquire introspective knowledge of his beliefs by perceiving the notebook. Gertler's argument relies on the claim that introspection is *peculiar* in the sense that it is different from other ways of knowing about the world. In contrast, my argument relies on the claim that introspection is *privileged* in the sense that it is epistemically more secure than other ways of knowing about the world. However, the relevant notion of epistemic privilege needs to be spelled out with care.

I don't claim that Inga's access to her memory is more *reliable* than Otto's access to his notebook. It is true that Otto could go temporarily blind or he could suffer from occasional visual illusions or hallucinations that lead him to misread the contents of the notebook. These cases must be the exception, rather than the rule, since the contents of the notebook do not constitute Otto's beliefs unless he has sufficiently reliable access to them. By the same token, however, Inga could suffer problems that prevent her from correctly retrieving information that is stored in memory. As Clark and Chalmers write:

Otto's access to the notebook is very reliable – not perfectly reliable, to be sure, but then neither is Inga's access to her memory. A surgeon might tamper with her brain, or more mundanely, she might have too much to drink. (1998: 15)

So we can stipulate that there is approximate parity between Inga and Otto with respect to reliability of access.

In my view, the epistemic difference between Inga and Otto is best described in terms of *rationality*, rather than *reliability*. Even if we stipulate that Inga and Otto are equally reliable, we do not thereby ensure that they are equally rational. And this is because Otto's misperceptions of the notebook need not constitute any failure of rationality, whereas Inga's memory lapses do. Let me explain.

First, why should we think that Inga's failure to access the information stored in memory constitutes a failure of rationality? Suppose Inga has the information that MoMA is on 53rd St stored in memory, although she fails to access the information on this occasion because of an error in retrieval. And suppose that Inga tries to answer the question whether she believes that MoMA is on 53rd St by using the procedure described below by Gareth Evans:

I get myself in a position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*. (Evans 1982: 225)

Using Evans' procedure leads her to believe that she doesn't believe MoMA is on 53rd St. Now she is in the Moorean predicament of believing that MoMA is on 53rd St, while believing that she doesn't believe this. But this is an irrational predicament. After all, if she conjoins these beliefs, then she believes the Moorean conjunction that *MoMA is on 53rd St and I don't believe it is*. Moreover, she cannot avoid rational criticism just by failing to conjoin her beliefs. If it's not rational to believe a Moorean conjunction, then it's not rational to believe the conjuncts of Moorean conjunction. To be sure, believing a Moorean conjunction is more egregiously irrational than believing each of its conjuncts, just as believing an explicit contradiction is more egregiously irrational than believing each of its conjuncts. But believing the conjuncts of a Moorean conjunction is no more immune from rational criticism than believing the conjuncts of an explicit contradiction.¹²

Second, why not think that Otto's failure to access the information stored in his notebook constitutes a failure of rationality? Suppose Otto has the information that MoMA is on 53rd St stored in his notebook, although he fails to access the information on this occasion because he has a visual illusion in which the notebook appears to say that MoMA is on 51st St. Using Evans' procedure leads him to believe that he does not believe that MoMA is on 53rd St. According to the extended

¹² See Smithies 2016 for a more extended discussion of Moore's paradox and its implications for an account of our introspective knowledge of standing belief.

cognition thesis, he is now in the Moorean predicament of believing that MoMA is on 53rd St, while believing that he doesn't believe this. So the extended cognition thesis predicts that Otto is irrational. But it's implausible to suppose that Otto is irrational just by virtue of misperceiving his notebook. Compare Stewart Cohen's (1984) new evil demon problem for reliabilism: a victim of a Cartesian evil demon can be fully rational even if their perceptual experiences systematically misrepresent the environment. Similarly, Otto can be fully rational even if his perceptual experiences occasionally misrepresent the contents of his notebook. In general, perceptual error is not sufficient for rational error.

In summary, a proponent of the extended cognition thesis cannot respect both of the following plausible principles about rationality:

- (1) *Pryor's principle*: if you're rational, and it perceptually seems to you that p , and there are no defeaters, then you believe that p . (Pryor 2000: 532)
- (2) *Shoemaker's principle*: if you're rational, and you believe that p , and you consider whether you believe that p , then you believe that you believe that p . (Shoemaker 1996: 83, 241-2)

These principles conflict in the case that Otto misperceives his notebook as saying that MoMA is on 51st St. According to Pryor's principle, it's rational for Otto to believe that his notebook says MoMA is on 51st St, rather than 53rd St. If he's rational, then Otto can use Evans' procedure in coming to rationally believe that he believes that MoMA is on 51st St, rather than 53rd St. And yet, by hypothesis, he believes that MoMA is on 53rd St. According to Shoemaker's principle, it's not rational for Otto to believe that MoMA is on 53rd St while believing that he does not believe this. So the two principles come into conflict.

A proponent of the extended cognition thesis might argue for epistemic parity between Inga and Otto by appealing to a perceptual model of introspection. On a perceptual model of introspection, there is no relevant epistemological disparity between Inga's introspective access to her memory and Otto's perceptual access to his notebook. However, the perceptual model of introspection should be

rejected precisely because it obscures the epistemological differences between perception and introspection. Here is an argument to that effect:

- (1) If the perceptual model of introspection is true, then introspection is vulnerable to *brute error*: there can be introspectively justified false beliefs about one's beliefs.
- (2) But introspection is immune from brute error: there cannot be introspectively justified false beliefs about one's beliefs.
- (3) Therefore, the perceptual model of introspection is false.

The rationale for premise (1) is that perception is vulnerable to brute errors in which one forms justified false beliefs on the basis of perceptual representations that misrepresent the world. The rationale for premise (2) is that if introspection is vulnerable to brute error, then there can be justified Moorean beliefs of the form, *p and I don't believe that p*. Since there cannot be justified Moorean beliefs of this form, there cannot be brute error in the case of introspection, and so the perceptual model of introspection is false.¹³

In conclusion, there is an important epistemic disparity between Inga and Otto. Inga's introspective access to her memory is immune from brute error, whereas Otto's perceptual access to his notebook is vulnerable to brute error. The claim is not that Inga's introspective access to her memory is more reliable than Otto's perceptual access to his notebook. There can be errors in memory retrieval just as there can be errors in perceptual representation. The difference is that errors in memory retrieval constitute failures of rationality, whereas errors in perceptual representation do not. The extended cognition thesis obscures this epistemic disparity between perception and introspection.

¹³ See also Shoemaker 1996: Ch. 2 & 11 and Burge 1996 for arguments against the perceptual model of introspection. The term 'brute error' is from Burge 1996, although I define it slightly differently from him.

7. The Argument from Disparity

In sections 1 and 2, we saw that there are plausible arguments for the extended cognition thesis and for accessibilism. In section 3, however, we saw an argument that these two positions are incompatible, and I defended the premises of this argument in sections 4-6. If the argument for incompatibilism is sound, then we cannot accept both accessibilism and the extended cognition thesis. Since I am persuaded by the Moorean argument for accessibilism, I am thereby committed to rejecting the extended cognition thesis. The aim of this section is to discharge the burden of explaining where the argument from parity goes wrong.

I claim that the third premise of the parity argument is false. This premise states that the information stored in Otto's notebook plays the same role as the information stored in Inga's memory. But in fact there is a crucial disparity between Inga and Otto – namely, that Inga has access to the information stored in her memory through introspection, whereas Otto has access to the information stored in his notebook through perception.

Moreover, this is not a merely superficial difference. On the contrary, it is deep enough to ground a cognitive difference between Inga and Otto. I've argued that the functional role of belief includes its role in justifying other beliefs. And I've argued that if access internalism is true, then beliefs can play this justifying role only if they are accessible through introspection. Since the information stored in Inga's memory is accessible through introspection, it can play the justifying role that is required for belief. In contrast, the information stored in Otto's notebook is accessible through perception, rather than introspection, so it cannot play the justifying role required for constituting belief.

The upshot is that we can give an *argument from disparity* for the conclusion that Otto's notebook, unlike Inga's memory, does not constitute beliefs:

- (1) Everything that plays a justifying role is accessible through introspection.
- (2) All beliefs play a justifying role.
- (3) The information in Otto's notebook is not accessible through introspection.
- (4) Therefore, Otto does not believe the information in his notebook.

As we have seen, proponents of access externalism reject premise (1), while proponents of strong phenomenal mentalism reject premise (2), and proponents of extended introspection reject premise (3). But having defended all three premises against objections, I conclude that we should accept the conclusion (4).

The parity argument fails, and the disparity argument succeeds, because extended states do not satisfy the parity principle. According to the parity principle, anything that plays the same role as belief is itself a belief. This principle is motivated by a broadly functionalist conception of the mind on which belief is realized by information that plays the right kind of role in the subject's psychology. But what exactly counts as the right kind of role? Different versions of functionalism give different answers to this question.

Clark and Chalmers emphasize the causal-explanatory role of belief. They argue that extended states can function much like beliefs in the causal explanation of behavior. On this basis, they conclude that extended states are beliefs. But their argument assumes that the role of belief is exhausted by its causal-explanatory role. Here is a representative passage from a recent discussion by Chalmers:

The deeper point is that extended states can function in explanation in very much the same way that beliefs function, and should be regarded as sharing a deep and important explanatory kind with them. This explanatory unification is the real underlying point of the extended mind thesis. (Chalmers 2011: xiv)

A key point that goes missing in Clark and Chalmers's discussion is that the functional role of belief is not exhausted by its causal role in explaining behavior, but also includes its epistemic role in justifying other beliefs. Given access internalism, beliefs must be accessible through introspection in order to play this justifying role. But since Otto's extended states are not accessible through introspection, they cannot play a justifying role. Therefore, Otto's extended states should not be counted as beliefs, even if they play an explanatory role that is otherwise similar to the explanatory role of belief.

The more general moral is that commonsense functionalism about belief is best construed as *normative* functionalism, rather than merely *causal* functionalism. In other words, belief should be defined in terms of normative as well as merely causal aspects of its functional role. After all, commonsense psychology is not a purely causal theory. It is also a normative theory of what justifies what. This is one way in which our commonsense theory of mind is different from other scientific theories. A functionalist theory of belief that highlights its causal role at the expense of its normative role thereby ignores what is most distinctive to our commonsense understanding (as opposed to a purely scientific understanding) of the mind.

Why have so many philosophers emphasized the causal role of belief at the expense of its normative role in developing functionalist theories of mind? Much of the impetus comes from a program of naturalistic reduction. Causation is thought to be more amenable to reduction than normativity and so better suited to giving a reductionist definition of mental states. But if the program of naturalistic reduction is to be successful, then it needs to encompass normativity as well as causation. So the program of naturalistic reduction provides no principled rationale for defining beliefs and other mental states in terms of purely causal rather than normative aspects of their functional role.

I conclude that the extended cognition thesis should be rejected on broadly epistemological grounds. And I put this forward as just one example of the way in which epistemology can and should constrain our understanding of the mind.

8. Conclusions

In this chapter, I have argued that we should reject the extended cognition thesis on epistemological grounds. I close with some clarifications and concessions.

First, I have argued that Otto's extended states are not beliefs because they are not accessible through introspection and so they cannot play an epistemic role in justifying other beliefs. But this is consistent with Clark and Chalmers' claim that extended states play a role in the causal explanation of behavior that is otherwise similar to the role of belief. So we might concede that extended states can play some but not all of the functional roles that we associate with our concept of belief.

Second, my discussion of the extended cognition thesis is exclusively concerned with doxastic states and processes, as distinct from the subdoxastic states and processes that figure in computational explanations in cognitive science. My discussion has no implications for a version of the extended cognition thesis that concerns subdoxastic states and processes, since they are not subject to epistemic evaluation and they do not play an epistemic role in justifying belief.

Finally, I do not deny the possibility of science-fiction scenarios in which beliefs are realized outside the skull. What I deny is that beliefs are realized outside the skull just by playing the kind of role that Otto's notebook plays. Otto's notebook does not realize beliefs because its contents are accessible by perception, rather than introspection, and so they cannot play an epistemic role in justifying other beliefs. Introspective access usually coincides with intracranial location, but the two can come apart: in science fiction scenarios, there can be introspective access to extracranial states and perceptual access to intracranial states. To that extent, I am in full agreement with Clark and Chalmers that "when it comes to belief, there is nothing sacred about skull and skin" (1998: 14).

Acknowledgements

This material was presented at the First Extended Knowledge Workshop at the University of Edinburgh in 2013 and as a keynote lecture at the Notre Dame/Northwestern Graduate Epistemology Conference and the British Philosophy Postgraduate Association Masterclass in 2015. Thanks to audiences on those occasions, especially Duncan Pritchard and Andy Clark, and to Adam Carter, David Chalmers, Brie Gertler, and an anonymous referee for helpful comments on a draft of this chapter.

References

- Adams, F. and K. Aizawa. 2001. The Bounds of Cognition. *Philosophical Psychology* 14: 43-64.
- Burge, T. 1986. Individualism and Psychology. *Philosophical Review* 95.1: 3-45.

- Burge, T. 1996. Our Entitlement to Self-Knowledge. *Proceedings of the Aristotelian Society* 96: 91-116.
- Carter, A., J. Kallestrup, O. Palermos, and D. Pritchard. 2014. Varieties of Externalism. *Philosophical Issues* 24.1: 63-109.
- Carter, A. & O. Palermos. 2015. Active Externalism and Epistemic Internalism. *Erkenntnis* 80.4: 753-72.
- Chalmers, D. 2011. Foreword. In A. Clark, *Supersizing the Mind*. Oxford: Oxford University Press.
- Clark, A. & D. Chalmers. 1998. The Extended Mind. *Analysis* 58.1: 7-19.
- Clark, A. 2010. *Memento's Revenge: The Extended Mind, Extended*. In R. Menary, *The Extended Mind*. Cambridge: MIT Press.
- Clark, A. 2011. *Supersizing the Mind*. Oxford: Oxford University Press.
- Cohen, S. 1984. Justification and Truth. *Philosophical Studies* 46.3: 279-95.
- Conee, E. and R. Feldman. 2001. Internalism Defended. *American Philosophical Quarterly* 38.1: 1-18.
- Conee, E. and R. Feldman. 2008. Evidence. In Q. Smith, *Epistemology: New Essays*. New York: Oxford University Press.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.
- Farkas, K. 2012. Two Versions of the Extended Mind Thesis. *Philosophia* 40: 435-47.
- Feldman, R. 2004. Having Evidence. In E. Conee, E. and R. Feldman, *Evidentialism: Essays in Epistemology*. New York: Oxford University Press.
- Gertler, B. 2007. Overextending the Mind. In B. Gertler and L. Shapiro, *Arguing About the Mind*. London: Routledge.
- Gibbons, J. 2006. Access Externalism. *Mind* 115: 19-39.
- Hurley, S. 1998. *Consciousness in Action*. Cambridge: Harvard University Press.
- Kratzer, A. 1981. The Notional Category of Modality. In H. Eikmeyer and H. Rieser, *Words, Worlds, and Contexts*. Berlin: De Gruyter.
- Pearlberg, D. and T. Schroeder. 2015. Reasons, Causes, and the Extended Mind. *Erkenntnis*.
- Pryor, J. 2000. The Skeptic and the Dogmatist. *Nous* 34.4: 517-49.

- Rupert, R. 2004. Challenges to the Hypothesis of Extended Cognition. *Journal of Philosophy* 101: 389-428.
- Schwitzgebel, Eric. 2002. A Phenomenal, Dispositional Account of Belief. *Nous* 36: 249-75.
- Shoemaker, S. 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Smithies, D. 2012. Moore's Paradox and the Accessibility of Justification. *Philosophy and Phenomenological Research* 85.2: 273-300.
- Smithies, D. 2014. The Phenomenal Basis of Epistemic Justification. In M. Sprevak and J. Kallestrup, *New Waves in Philosophy of Mind*. London: Palgrave Macmillan.
- Smithies, D. 2016. Moore's Paradox, Belief, and Self-Knowledge. *Philosophical Issues* 26: XX-XX.
- Sprevak, M. 2009. Extended Cognition and Functionalism. *Journal of Philosophy* 106: 503-27.
- Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT Press.
- Weiskopf, D. 2008. Patrolling the Mind's Boundaries. *Erkenntnis* 68.2: 265-76.