Minds, materials and metaphors

ADAM TOON

'the mind extends into the world'
(Clark and Chalmers, 'The Extended Mind')

'the mind is not even a metaphorical "place"'
(Ryle, *The Concept of Mind*)

Abstract

What is the relationship between mental states and items of material culture, like notebooks, maps or lists? The extended mind thesis (ExM) offers an influential and controversial answer to this question. According to ExM, items of material culture can form part of the material basis for our mental states. Although ExM offers a radical view of the *location* of mental states, it fits comfortably with a traditional, representationalist account of the *nature* of those states. I argue that proponents of ExM would do better to adopt a *fictionalist* approach to mental states. In so doing, I suggest, they could retain the important insights underlying the extended mind thesis, while avoiding its more problematic consequences.

1 Introduction

What is the relationship between mental states and items of material culture, like notebooks, maps or lists? The extended mind thesis (ExM) offers an influential and controversial answer to this question. According to ExM, our mental states can, under the right circumstances, be partly constituted by items of material culture: together with our brains, things like notebooks, maps or lists can form part of the physical stuff that realises our beliefs and desires. In this way, the mind can extend beyond the brain and body into the world. ExM is often seen as a radical view and has accordingly sparked a good deal of debate (for an overview, see Menary, 2010). And yet, while ExM might offer a radical view of the *location* of mental states, it fits comfortably with a traditional, representationalist account of the *nature* of those states (Crane, 2016). Indeed, in its canonical form due to Clark and Chalmers (1998), ExM begins with the assumption that beliefs are representational states with a particular sort of causal role. On top of this basic picture, ExM simply adds the idea that these states can sometimes extend outside the head.

This paper will offer an alternative analysis of the relationship between mental states and material culture. In particular, I will

suggest that we adopt a *fictionalist*, rather than representationalist, account of the nature of mental states. According to the fictionalist, our ordinary talk about the mind is fundamentally metaphorical: we treat people *as if* they had inner representations that express the content of their mental states, but we don't take this too seriously. Recently, a number of authors have begun to explore fictionalist approaches to folk psychology (e.g. Demeter 2013; Toon, 2016; Wallace 2007, 2016). In this paper, my aim is not to offer general arguments in favour of fictionalism, or against representationalism. Instead, my aim is to show that, by adopting a fictionalist approach to the mind, we can provide a novel and compelling account of the relationship between mental states and material culture. It is not that mental states are typically inner representations that sometimes, under special circumstances, extend *outwards* into the world. Instead, much of our talk about the mind is itself a metaphorical projection *inwards* from the world of material culture. And yet, I shall argue, the fictionalist can still make sense of the notion of extended mental states. The result is a view that steers a middle course between the two quotes that began this paper. *Contra* Ryle's remark, the mind *is* a metaphorical place – a metaphorical world of inner sentences and pictures. But precisely because this place is metaphorical, and not literal, the idea that it can extend into the world must be handled with care.

Although it departs from the usual way of understanding the notion of extended mental states, a fictionalist approach echoes important themes from the wider literature on the interactions between minds, tools and material culture. Especially important here is the work of Daniel Dennett (e.g. 1996). Although he has rejected the label of 'fictionalist', the approach I shall adopt shares some important similarities with Dennett's views on folk psychology. While Dennett's writing on tools is often cited in discussion of the extended mind thesis, his account of the nature of mental states has figured far less in the debate. Interestingly, Clark himself has expressed sympathy for Dennett's views on folk psychology and mental states, especially in his work on connectionism (Clark, 1989; 1993). If the argument in this paper is along the right lines, a fictionalist approach to the mind may provide a better way to accommodate the key insights behind the extended mind thesis, while avoiding some of the most important objections levelled against it.

The discussion will proceed as follows. First, I will introduce the extended mind thesis and consider its relationship to representationalism (Section 2). I will then introduce mental fictionalism (Section 3) and show how it offers a new perspective on the relationship

182

between mind and material culture (Sections 4 and 5). Finally, I will show how this approach allows us to overcome a number of well-known objections to ExM (Sections 6 to 9).

2 ExM and representationalism

Clark and Chalmers (1998) introduce the idea of extended mental states through the famous example of Otto and Inga. Inga hears of an exhibition at the Museum of Modern Art (MoMA). She recalls that the museum is on 53rd Street and heads off to see it. In doing so, Inga relies upon her ordinary, biological memory and not upon any external devices. By contrast, Otto is an Alzheimer's patient who suffers from memory loss and therefore carries a notebook with him at all times to record important information. When Otto hears about the exhibition, he looks up the information in his notebook and heads off to see it. Clark and Chalmers argue that, despite their apparent differences, Otto's notebook plays a similar role in his life to the role that Inga's biological memory plays in hers. As a result, they claim, Otto's beliefs are partly constituted by the entries in his notebook. Otto believes that the exhibition is on 53rd Street even before he looks at his notebook, in much the same way that Inga believes this even before consulting her biological memory. The result is that Otto's mind extends beyond his head and body and into the world.

At the heart of Clark and Chalmers' argument is the *parity principle*:

> If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim) part of the cognitive process (1998, p. 8).

The parity principle is intended as a heuristic device, encouraging us to form judgments about what counts as cognition and mind behind a 'veil of metabolic ignorance' (Clark, 2008, p. 114). In the case of Otto, it asks us to imagine what we would say if the information about MoMA were located, not in the pages of a notebook, but inside his head. Would we count it among Otto's beliefs? If so, Clark and Chalmers argue, then we ought to say the same about the entry in his notebook. After all, the only difference between the two is the whereabouts of the relevant information. Of course, this does not mean that every time we use a notebook, we acquire an extended belief. Clark and Chalmers stress that Otto's notebook satisfies

183

conditions of 'glue and trust' (Clark, 2010b, p. 83): Otto always has the notebook with him, trusts what it says as a matter of course, and where its information might be relevant, he rarely acts without consulting it (Clark and Chalmers, 1998, p. 17). Most uses of notebooks in everyday life will fail to meet one or more of these conditions. If these conditions are met, however, Clark and Chalmers argue that we have a case of extended belief.

The extended mind thesis has provoked a range of objections. Many have seen it as a radical view that runs counter to our ordinary talk about the mind. Others have argued that it overlooks important differences between biological memory and items of material culture. I shall return to these objections later on. For now, however, it is important to notice that, despite the controversy it has caused, Clark and Chalmers' argument for ExM fits comfortably with a traditional, representationalist conception of the nature of belief (Crane, 2016). To simplify somewhat, representationalism holds that having a belief involves having a mental representation which expresses the content of that belief (Schwitzgebel, 2015). For example, to believe that Derby County will win promotion is to have a mental representation with the content *Derby County will win promotion*. Other mental states are understood in a similar manner. To desire that Derby will win promotion – or to hope or doubt or fear that they will – each involves possessing a mental representation with the same content. The difference between these mental states is normally understood in terms of the different causal roles that the representation plays in the production of behaviour. Thus, the desire that Derby will win promotion might cause someone to cheer loudly in support, while the fear that they will might lead them to cheer for their opponents instead. Despite some notable challenges, it seems fair to say that representationalism remains the dominant view of the nature of mental states.

As presented by Clark and Chalmers, the extended mind thesis is entirely compatible with this understanding of the nature of mental states. Indeed, their discussion begins by adopting a broadly representationalist conception of belief in order to argue for the notion of extended mental states: the argument proceeds by pointing to the entries in Otto's notebook and arguing that, since they play a similar role to Inga's biological memory, they ought to count as bearing the content of Otto's beliefs. As Clark and Chalmers put it, '[t]he information in the notebook functions just like the information *constituting* an ordinary non-occurrent belief; it just happens that this information lies beyond the skin' (1998, p. 13; emphasis added). What the extended mind thesis adds to the standard representationalist story is the idea that the representations that bear the content

184

of our mental states can sometimes be found outside the head. While this addition has certainly proved controversial, it leaves the basic tenets of representationalism intact. In what follows, I will suggest that proponents of ExM would do better to abandon representationalism and turn to fictionalism. Let us now consider the main features of a fictionalist approach to mental states.

3 Mind as metaphor

Fictionalism claims that our ordinary talk about the mind is fundamentally metaphorical. According to the fictionalist, we treat people *as if* they had inner representations that express their beliefs or desires, but we don't claim that they actually have such things inside their heads. One way to fill out this idea is to draw on an influential analysis of metaphor and figurative language due to Kendall Walton (1993). Suppose that someone asks us where the town of Crotone is and we reply, 'It's on the arch of the Italian boot'. Walton suggests that we understand such talk in terms of pretence within a game of make-believe. When we talk about Italy in this way, we are invoking a familiar game of make-believe in which we imagine Italy to be a boot. Our utterance is an act of pretence in this game. We do not really claim that Crotone lies on the arch of a giant boot floating in the Mediterranean; we merely pretend to assert this. And yet, by doing so, we indicate that pretending in this way is appropriate within the game. If we had said, 'Crotone is on the shoelaces of the Italian boot', our pretence would have been inappropriate. Why? Because Crotone lies in the southern coast of Italy, somewhere between Capo Colonna and Taranto. *This* is what we actually assert with our utterance. Pretending that Italy is a boot provides us with a vivid and convenient way of making this assertion.

Fictionalism takes a similar approach to talk about the mind as an inner world. We can introduce this approach using a twist on Wilfred Sellars' famous myth about the origin of talk about mental states (Sellars, 1956). Sellars imagines a society that uses a 'Rylean' language that refers only to overt behaviour. Into this society comes a visionary called Jones, who develops a theory of internal, psychological episodes he dubs *thoughts*. Jones bases his theory of thoughts on overt verbal behaviour. Thoughts are said to be similar to spoken utterances in some respects (e.g. they have content) but dissimilar in others (e.g. they are not uttered by any inner tongue). Using his theory, Jones can explain and predict people's behaviour in a way that the original Rylean language could not. In the fictionalist

185

Adam Toon

version of this myth, Jones is not a visionary theorist. Instead, he is rather like the first person to describe Italy as a boot. What Jones adds to the Rylean language is not a new theory of an inner realm, but a useful metaphor. He invents a game of pretence in which we talk about people *as if* they undergo inner episodes that are analogous to overt verbal behaviour. This helps us to make sense of people and their behaviour, just as talking as if Italy were a boot helps us to make sense of its geography.

Does fictionalism deny the existence of mental states? Consider another case of metaphor. Suppose we say 'the clouds are angry today' (cf. Walton, 1993). For the fictionalist, the question 'do mental states exist?' is like the question 'do angry clouds exist?' In both cases, the answer is: in one sense 'yes' and in another sense 'no'. Are there clouds that are literally angry (or happy or miserable)? No. Are there clouds that are quite properly called angry (or happy or miserable) when we're indulging in a particular sort of pretence? Yes. The same lesson applies to talk about mental states. Are there literally representations inside people's heads that express their thoughts (or beliefs or desires)? No. Are there 'real patterns' (Dennett, 1991) in people's behaviour that are picked out when we invoke this pretence? Yes. It would be misleading, then, to say that fictionalism denies the existence of mental states. *If* mental states are presumed to be inner representations, then fictionalism does indeed deny that such things exist. Our talk about these inner representations is only metaphorical. But we use these metaphors to pick out features of people's behaviour that are perfectly real and non-metaphorical. In this sense, fictionalism does not deny the existence of mental states. Instead, it offers us an account of what mental states are and how they are picked out by folk talk about the mind.

Elsewhere I have argued that this view of mental states has much to recommend it over well-known alternative approaches to the mind (Toon, 2016). For example, unlike analytical behaviourism, fictionalism does not try to reduce talk about the mind to talk about behaviour. Indeed, it explains why this project is doomed to failure: like many metaphors, those we use to talk about the mind allow us to express things that cannot be given a literal paraphrase. Unlike instrumentalism, fictionalism can acknowledge the intuition that talk about the mind does, in a sense, involve the idea of an inner realm – it is simply that this inner realm is metaphorical, not literal. And unlike eliminative materialism, fictionalism allows us to retain our ordinary talk about belief and desire, whatever future cognitive science might uncover about what happens inside our heads. There are, of course, other approaches to the mind that also depart from representationalism, notably enactivism

186

(e.g. Varela, Thompson and Rosch, 1991; Hutto and Myin, 2013). My aim, however, is not to defend fictionalism against each of these competing approaches, but to show that it offers us a fresh perspective on the relationship between mind and material culture. Let us now turn to consider this aspect of fictionalism in detail.

4 Materials as metaphors

In our re-telling of Sellars' myth, the source of Jones' metaphor for thoughts was overt verbal behaviour. Other mental phenomena are better served by other metaphors, however. The world of public representations is enormously varied: it contains many different forms of representation used in many different ways. As a result, it offers a vast stock of metaphors for talking about the mind. To take an obvious example, overt verbal behaviour is typically fleeting. A sentence is uttered in one moment and lost in the wind the next. By contrast, written language is longer lasting and is put to different uses as a result. This also means that it lends itself to different metaphorical purposes in our talk about mental states.

Consider memory and standing beliefs. From the perspective of the fictionalist, what is so striking about Otto is that he offers a vivid description of the source of one of our central metaphors for talking about the mind. Thinking of memory as if it were a trusty, ever-present notebook lodged away somewhere inside our heads is a pervasive and powerful way of making sense of people and their behaviour. When someone encounters some important fact or situation in the world, we can imagine them jotting it down in their inner notebook. Later, we can imagine them looking up the information that they've written down and using it to guide their actions. Indeed, the rules that we suppose to govern the use of this inner notebook are summed up fairly well in Clark and Chalmers' description of Otto, especially the 'glue and trust' conditions: our inner notebook is always present, fairly trustworthy and we always consult it when we need it. In this way, the fictionalist suggests, we draw upon our understanding of our practices for using certain forms of external representation – in this case, notebooks – as the source of a metaphor for making sense of people and their behaviour. (Of course, these metaphors can also mislead us. We shall return to this point in Section 9. For a history of the different metaphors that have been applied to memory and their pitfalls, see Draaisma, 2000.)

The metaphor of memory as a notebook acts to fill out one region of our folk psychological discourse. Looking at Italy as if it were a boot

yields the rules of the Italy-as-a-boot game, allowing us to make as-sertions about the coastline of Italy by making utterances within this game. In a similar manner, looking at people as if they possessed an inner notebook yields the rules of the game for attributing stand-ing beliefs. If we say 'Inga believes that MoMA is on 53rd Street', we are not claiming that she has a representation with this content lodged somewhere in her head; we are indicating how to pretend correctly within our game. If we had said, 'Inga believes that MoMA is on 52nd Street', our pretence would have been inappropriate. Why? Because of a whole range of facts about Inga's behaviour: when she hears about the exhibition, she heads off to 53rd Street, not 52nd Street; when someone asks her for directions to MoMA, she says '53rd Street', not '52nd Street'; and so on. The behaviour that makes our pretence appropriate will, more often than not, be enor-mously varied. This is one reason why talk about mental states cannot be reduced to talk about behaviour. Nevertheless, it is still facts about Inga's behaviour that make our pretence appropriate (or inappropriate). The metaphor of memory as a notebook gives us a way of picking out this complex pattern in Inga's behaviour, just as the metaphor of Italy as a boot gives us a way of describing its coast-line. Put simply, Inga behaves *as if* she had 'MoMA is on 53rd Street' written in her inner notebook – even though there is no such notebook.

Notebooks are not the only sources of metaphors that we use for making sense of the mind. Other forms of external representations provide important inspiration too. Think of an architect's drawings (cf. Houghton, 1997). We call such drawings 'plans', of course. And it is precisely external, material representations like these – along with other documents that fulfil similar functions, like route maps for upcoming journeys or itineraries for day trips – that provide our means for talking about what people do when trying to achieve certain goals. Suppose that Ruth wants to put up some shelves in the living room. How might she go about it? She might write down a list of what she needs to do (go to the shed, fetch the tool box, mark out screw holes, etc.) and draw a sketch of how she wants to position the shelves. Alternatively, she might manage without pencil and paper. Even if she does, however, we can make sense of her behaviour (her trip to the shed, her drawing pencil marks on the wall in such-and-such a way, etc.) by imagining that she did have such a list and sketch guiding her actions. Similarly, we can think of someone's desires as if they were a kind of private, inner shopping list (or 'wish list') that they will try to tick off when they get the opportunity.

188

In Section 2 we saw that the extended mind thesis begins from a representationalist account of belief. Beliefs are taken to be mental representations with particular causal roles. What ExM adds to this basic vision of the mind is the idea that these representations can sometimes be said to extend outwards into the world. In a sense, the fictionalist approach is precisely the reverse. For the fictionalist, the mind itself is a metaphorical projection *inwards* from the world of material culture. In many respects, it is external representations that come first. In particular, we draw on cases in which people use external representations to talk about cases in which they do not. The question remains, however, of what the fictionalist should say about cases in which people actually *do* rely upon external representations, like notebooks, blueprints and shopping lists? Can fictionalism make sense of the idea of extended mental states? My aim in the rest of this paper will be to show that it can and that, in fact, fictionalism provides an alternative way to understand ExM that can avoid many of the objections levelled against it.

5 Materials as minds

Can fictionalism make sense of the idea of extended mental states? For the fictionalist, terms like 'belief' and 'desire' are fundamentally metaphorical. So our question becomes: in what cases should we apply these metaphors? In what cases are such metaphors apt or useful? In particular, should we apply them only in cases where people rely on their brains and bodies alone – or are they also useful when people are dependent on external devices? At first glance, this might seem like an odd question. After all, the fictionalist claims that the world of external representations is the *source* of our metaphors for describing the mind. Using the world of external representations as a metaphor for talking about the world of external representations would not seem terribly helpful: we seem to be asked to apply the metaphor to its own source. Asking someone to see Italy as a boot might be a helpful way of describing its coastline, but asking them to see a boot as a boot is not especially illuminating.

When we look more closely, however, we see that fictionalism is able to make sense of the notion of extended mental states. For often we use one sort of external representation to make sense of another. Consider memory and standing belief. According to the fictionalist, our talk about memory is guided by a core metaphor of memory as an ever-present and trusted notebook. We use this metaphor to make sense of how people behave without actual notebooks to

189

hand. And yet we might also apply it when people rely on other tools and external representations. Suppose that, rather than writing down addresses, Otto drew little maps and sketches in his notebook to help him find his way around. We might still say that Otto believes that MoMA is on 53rd Street, even if this is not written down anywhere in his notebook. What we are doing in this case, according to the fictionalist, is using our understanding of one form of representation to make sense of another. We are saying that, for some purposes, we won't go too far wrong if we treat Otto as if he had a notebook with the address written down. Saying this does not depend upon controversial claims about being able to reduce the content of maps or pictures to propositions. It simply relies on the fact that people can *do* similar things with maps and pictures as they can with written text.

The picture that emerges is as follows. Humans have a remarkable capacity to use all manner of different external, material representations. This is an enormously important feature of human life in its own right, of course. According to the fictionalist, it also shapes our vision of the human mind as an inner realm. For some of the most important ways in which we use external representations – for recording information, stating our wishes, for planning our actions – also lie at the heart of our language for talking about the mind. We use these external devices as the source of metaphors for making sense of people and their behaviour. It is these metaphors that give meaning to our attributions of mental states like belief, desire and intention. And yet when we come to apply these metaphors, we do not restrict ourselves to human activity in the absence of tools and external representations. When we attribute states like belief and desire, our main concern is to make sense of people's behaviour. Often, that behaviour is itself dependent upon interaction with certain sorts of tools and external representations – sometimes ones that are quite different from the original source of our metaphors.

Fictionalism yields a rather different set of criteria for attributing extended mental states to the usual, representationalist interpretation of ExM. Recall that, according to representationalism, beliefs are mental representations that play a characteristic causal role. When we attribute a belief to someone, we are claiming that their cognitive machinery contains such a representation. Clark and Chalmers' strategy is to show that, in some cases – such as Otto's – representations that play the appropriate causal role can be found outside the head. On this view, to judge whether we are dealing with a genuine case of an extended mental state, we should ask ourselves: does this external representation play the right causal role to count as a belief (or

190

desire, or intention, etc.)? Since our intuitions are taken to be clearest when applied to inner states, the parity principle suggests that we imagine that the external representation was instead found internally before delivering our judgement.

In contrast, fictionalism suggests that even attributions of ordinary, non-extended mental states involve an imaginative move akin to that proposed by the parity principle: we make sense of people by imagining that they had inner analogues of external representations like notebooks (or lists or itineraries etc.). When judging whether we are faced with a case of an *extended* mental state, we should no longer focus on a particular external representation and ask ourselves whether it plays the appropriate causal role to count as a belief. Instead, we should ask: does the person using this external device display the appropriate overall pattern of behaviour for us to apply our stock of folk psychological metaphors? And when we answer this question we rely not upon a prior grasp of beliefs as inner states, but on our understanding of the way people use external, public representations, like notebooks.

Let us now see how reinterpreting the notion of extended mental states in this way can help us to respond to some well-known objections to ExM. In each case, I will suggest, the root of the problem lies not with ExM but with representationalism – and turning instead to fictionalism shows us a way out of trouble.

6 Common sense

The first, and most straightforward, objection to the extended mind thesis that we will consider is simply that it clashes with our common sense view of the mind. Clark and Chalmers try to motivate ExM by appealing to our ordinary, folk psychological conception of belief. And yet, it is argued, we ordinarily think of mental states as inside the head. The folk are internalists. Talking of beliefs extending into notebooks therefore stands radically at odds with our common sense view of the mind (e.g. Adams and Aizawa, 2001; Rupert, 2004).

This objection is less troubling than it might seem at first sight. When we consider ordinary talk about the mind more closely, I think, we see that we often attribute extended mental states. David Houghton (1997) offers a number of examples. Suppose that, before heading out to the supermarket, Ted inspects his kitchen cupboards and writes down a list of items he wants to buy: a pint of milk, bread, cornflakes, carrots, and so on. It would be quite natural to say that Ted wants to buy, say, milk and carrots, even if he's unable to

191

recall these items without his list to hand. After all, this is precisely the reason that we use shopping lists. Or consider an architect's technical drawings. It would be entirely in line with our ordinary way to talking to say that these drawings capture the architect's intentions for the building. Saying this does not rely upon assuming that the architect must have committed every aspect of her drawings to memory. Indeed, if the drawings are sufficiently complex, it might not even be possible for the architect to do this. And yet we should still say that she intends to construct such-and-such a building.

So there are reasons to think that ExM might not depart so far from our ordinary talk about the mind after all. In fact, Houghton argues, it is internalism that is at odds with ordinary usage. Many remain unconvinced, however, even those who are otherwise sympathetic to ExM. Thus, Mike Wheeler imagines asking the folk a question that directly concerns the whereabouts of a particular cognitive state, such as 'Where in space are the relevant cognitive states of the architect realised?' (2011, p. 424). Not implausibly, Wheeler suggests that the likely response is an internalist one. So it seems that folk opinion is rather conflicted: on the one hand, we commonly attribute extended states; on the other hand, we baulk when asked about the location of those states.

Fictionalism can explain these apparently conflicting responses to ExM. As we have seen, the fictionalist claims that all attributions of mental states involve metaphor: we talk *as if* people were guided by representations that capture the content of those states. And it is the metaphorical nature of mental state attribution, I believe, that explains why our intuitions falter when asked explicitly about the whereabouts of mental states. For, properly speaking, mental states are nowhere. To ask where beliefs *are* is to ask a 'silly' question – one that pushes a metaphor too far (Yablo, 1998; cf. Di Paolo, 2009). It is rather like being told that someone has a chip on his shoulder and asking if it is on his right or left shoulder. Notice too that this feature of our ordinary talk about the mind is not confined to cases of extended mental states. To ask whether beliefs lie inside or outside the head is, I think, no more (or less) silly than asking whether they lie on the left or right side of the skull. Both questions misunderstand the nature of our talk about mental states. Both push a metaphor too far.

7 Intentionality

A rather different line of criticism levelled against ExM has been to allege that it ignores a crucial distinction between the intentionality

of mental states and that of public representations, such as spoken and written language. It is widely held that the intentionality of public representations is derived from the intentionality of mental states. Furthermore, according to representationalism, the intentionality of mental states is, in turn, to be explained in terms of the content of mental representations. In contrast to public representations, mental representations are said to possess *original* intentionality: their content does not depend upon any other intentional states. Instead, the content of mental representations is to be explained in non-intentional terms, such as causal relations or evolutionary history (e.g. Stich and Warfield, 1994). From this perspective, critics have argued that the entries in Otto's notebook cannot count as his beliefs, since they possess merely *derived*, rather than original, intentionality: their content depends upon social conventions and the mental states of those who use them (e.g. Adams and Aizawa, 2001, 2008; Fodor, 2009).

Fictionalism rejects this approach to intentionality. To say that someone has a certain mental state is not to say that they possess a particular inner representation; it is to say that they behave *as if* they had such a representation. For the fictionalist, then, the intentionality of mental states cannot be grounded in the intentionality of mental representations, since she argues that these inner representations do not exist. Instead, the intentionality of mental states is ultimately grounded in facts about a person's behaviour: it is because someone behaves in the way that they do that they can properly be said to possess certain mental states. Moreover, exhibiting the right pattern of behaviour to count as possessing a given mental state will typically involve engaging in the use of public representations, especially language. For example, amongst the behaviour that makes it appropriate to say that Inga believes that MoMA is on 53rd Street is that, if you ask her where MoMA is, she'll answer '53rd Street'. The upshot is that fictionalism claims that the intentionality of mental states is derived from the intentionality of public representations, rather than vice versa.

This is an important feature of mental fictionalism. One reason it is important is that it allows us to address a serious difficulty facing the approach, often called the problem of 'cognitive collapse' (e.g. Wallace, 2016). The fictionalist claims that inner representations do not exist and that our talk about them is merely a useful fiction. And yet a fiction is itself a representation: our folk psychological fiction, for example, represents people as having inner representations. To avoid collapsing like a house of cards, it seems that fictionalism must therefore allow that at least one representation exists and

193

has content. And yet, once we allow this much, why deny the existence of mental representations in particular? And how can we explain the existence of any sort of representation or content without them? The solution is to notice that fictionalism has no difficulty in granting the existence of external, public representations that are meaningful, such as written and spoken language. These public representations gain their meaning from their role in norm-governed social practices, not from representations lodged inside anyone's head (cf. Wittgenstein, 1953). Our folk psychological fiction is itself a form of public representation: it involves acts of pretence performed within a norm-governed social practice (a game). Why think that talk about inner representations is especially problematic? Because our ordinary concept of representation applies to objects used in certain social practices, like words, maps or diagrams. Talking as if the mind were an inner world of representations is enormously useful. And yet, when we stop to think about it, the idea that people could really have such things inside their heads makes little sense. (I offer a fuller articulation and defence of these claims in Toon, forthcoming.)

Not all will accept this approach to intentionality, of course. In the present context, the important point is that it allows ExM to avoid the objection that it ignores the distinction between the intentionality of mental states and public representations. According to the fictionalist, what happens in cases of extended mental states is *not* that public representations are somehow illegitimately co-opted to serve as mental representations. Instead, what is distinctive of such cases is that, by interacting with public representations or other external objects, someone manages to exhibit the right overall pattern of behaviour to count as possessing an extended mental state. Thus, by interacting with his notebook, Otto is able to exhibit broadly the same pattern of behaviour as Inga – he too can reply '53rd Street' when you ask him MoMA's address, for example. In this way, cases of extended mental states can be, at one and the same time, instances of *both* forms of intentionality: they are cases in which the intentionality of public representations and the intentionality of mental states can overlap (cf. Adams and Aizawa, 2008, p. 38).

Notice, however, that although these two forms of intentionality *can* overlap, they need not do so. In this respect, Otto's case is potentially misleading. In Otto's case, the content of his notebook entry and the content of his belief is the same. In other cases, however, they can come apart. To see this, notice that, for the fictionalist, Otto counts as having the belief that MoMA is on 53rd Street *not*, strictly speaking, in virtue of the fact that his notebook contains a

194

representation with this content; instead, it is in virtue of his overall pattern of behaviour. In principle, Otto might exhibit the same pattern of behaviour even if the content of his notebook entries were different. For example, suppose that Otto realises he has made a mistake when originally writing down addresses in Manhattan: he counted the street numbers wrongly and they're all one number too low. If he's feeling lazy, he might not bother updating his notebook but just remember to add one every time he looked them up. In this case, the content of his belief might remain the same (he believes that MoMA is on 53rd Street) while the content of his notebook entry is different (it now says 'MoMA is on 52nd Street').

More dramatically, we might even imagine cases in which an external device bears no representational content at all *qua* external object and yet still enables someone to possess extended mental states. Suppose that James suffers from anxiety, but finds that using a stress ball helps him to remain calm (cf. Colombetti and Roberts, 2015). Without his stress ball, let us assume, James finds himself rather overwhelmed at work: he fears that tasks will never be completed, worries that colleagues doubt his abilities, wishes he could avoid responsibilities, plans to find a new job, and so on. Thankfully, with his stress ball to hand, he finds matters altogether more manageable. As a result, we might attribute all manner of more positive beliefs, desires and intentions to James: he now believes that the troublesome matter with his client can be resolved, that his boss values his contribution to the team, wishes to handle the new contract himself once it is awarded, plans to build his career at the company, and so on. Like Otto's notebook, the stress ball allows James to possess extended mental states. But it achieves this feat not by expressing the content of those states, but by helping to regulate and guide James's overall behaviour in such a way that our attributions are appropriate.

8 Cognitive science

As well as questioning ExM's credentials vis-à-vis common sense, critics have also worried about its implications for cognitive science. The main concern here is that extended cognitive systems will simply be too heterogeneous to form a proper basis for scientific investigation. Thus, Adams and Aizawa ask us to:

> Consider […] the range of tools humans use as mnemonic aids. There are photo albums, Rolodexes, computer databases,

195

> strings around the finger, address books, sets of business cards, bulletin boards, date books, personal information managing software, palmtop computers, hand drawn maps, and lists of 'things to do'. What are the chances of there being interesting regularities that cover humans interacting with all sorts of tools? Slim to none, we speculate. There just isn't going to be a science covering the motley collection of 'memory' processes found in human tool use (2001, p. 61)

Clark (2010a) considers two main ways of responding to this worry. The first, more optimistic, response points out that even apparently dissimilar causal mechanisms might turn out to display regularities when seen from a suitable overarching framework, such as that of 'information storage, transformation and retrieval' (2010a, p. 50). The second, more pessimistic response that Clark considers (but does not endorse) takes the real lesson of ExM to be that 'the realm of the mental is itself too disunified to count as a scientific kind' (2010a, p. 62) and that perhaps we ought therefore to '*eliminate the mind*' (*ibid*., p. 63; emphasis in original).

Fictionalism allows us to avoid the threat of eliminativism for mental states. We seem to be faced with a dilemma: either terms like 'belief' and 'desire' pick out some underlying state or processes that form a scientific kind, or else they must be eliminated. But this is a false dilemma. After all, there are many other legitimate uses of language, even in scientific contexts. According to fictionalism, terms like 'belief' and 'desire' function as metaphors that allow us to pick out certain patterns in behaviour. What unites cases of extended belief is that they each exhibit this pattern of behaviour, not that they share any common inner state or process. As a result, it is no argument against ExM – at least if we adopt a fictionalist approach – to point out that cases of extended belief (or desire or intention etc.) might contain an unruly motley of different internal processes.

Here we might recall Dennett's well-known example of Jacques, the Frenchman who kills his uncle in Trafalgar Square (Dennett, 1987, p. 54). Jacques is caught by Sherlock, while Tom reads about the crime in the *Guardian* and Boris in *Pravda*. All four share a belief: they all believe that a Frenchman has committed murder in Trafalgar Square. And yet, Dennett submits, it is hard to see why they must all share a similarly structured object inside their heads. The fictionalist shares this intuition for ExM cases too. Consider James again, who uses a stress ball to alleviate anxiety. Now consider his colleagues Jess, John and Jane. Jess achieves a similar boost to her mood by relying on a book of self-help mantras, while John relies

196

instead on anti-depressants. Finally, Jane is more naturally relaxed and confident. All four might share the belief that the troublesome difficulty with their client will be resolved. And yet it is hard to see why they must all share a similarly structured object either inside or outside their heads.

So the fictionalist need not worry that cases of extended mental states are unlikely to show much unity in their underlying states or processes. At the same time, fictionalism can explain why we might be misled to expect such unity. For, according to the fictionalist, we make sense of *all* cases of belief – however diverse their underlying machinery, whether extended or not – by treating them *as if* they were guided by inner sentences that play a particular role. If we begin to take such talk too seriously, then we will be tempted to misconstrue terms like 'belief' and 'desire' as attempts to pick out real inner states or processes with distinctive characteristics, such as possessing a certain sort of content or playing a particular causal role. Seen correctly, however, our ordinary talk about the mind does not carry such implications.

Of course, even if the existence of a motley of underlying processes does not threaten the idea of extended mental states, we might still worry about its implications for cognitive science. Would any attempt to construct a science of such a diverse assortment of systems, states and processes be doomed at the outset? It is hard to see why. We must be careful not to hanker after an unrealistic ideal of the unity of science. After all, even the principles of Newtonian mechanics are, strictly speaking, true of few real systems. Instead, Newtonian mechanics offers a set of models (free fall, the ideal oscillator, and so on) that apply more-or-less accurately in a diverse range of areas, from falling leaves to fluid flows (e.g. Giere, 1988). It is not clear why a science of extended cognition might not aim for a similar unity. Moreover, even if this aim proves unrealistic in the long run and cognitive science turns out to be rather less unified than Newtonian mechanics, this would hardly be especially damning. Most sciences must invoke a range of different principles and models to encompass their subject matter (Clark, 2010a).

9 Drawing boundaries

We can now consider a final line of criticism concerning ExM. As well as ignoring a crucial distinction concerning intentionality, ExM has also been charged with overlooking important differences in the causal profiles of internal and external processes. In this vein, a number of critics have drawn attention to differences between

biological memory and external devices like Otto's notebook (e.g. Adams and Aizawa, 2001; Rupert, 2004; Weiskopf, 2008). For example, psychologists have found that ordinary, biological memory exhibits a phenomenon called *negative transfer*, whereby old memories interfere with the ability to form new ones. If I have memorised your old phone number, for instance, it will be harder for me to memorise your new number than it would be if I'd never known the old one in the first place. When using his notebook, it seems that Otto won't exhibit negative transfer: he can just rub out the old phone number and write down the new one. Similar points can be made for other important characteristics of biological memory that psychologists have described, such as so-called generation effects or power laws of remembering and forgetting (Rupert, 2004; Sprevak, 2009). Given such differences, critics argue, the entries in Otto's notebook do not count as beliefs since it is simply false to say that they play the same causal role as items in Inga's biological memory.

In response, proponents of ExM have justly questioned whether these features of biological memory are essential to belief. Recall that Clark and Chalmers' claim is that the notebook entries play the causal role *that folk psychology attributes to beliefs*. It is surely no part of the folk notion of belief that people's memory exhibits effects like negative transfer. After all, describing such effects has taken detailed scientific investigation. This point is often expressed in terms of the 'grain' in which we ought to specify the relevant causal roles. Effects like negative transfer are said to be too fine-grained to appear in the folk notion of belief. If we adopt a more coarse-grained specification of the causal role of belief, then Otto's notebook can count as part of his mind (Clark, 2010a; Sprevak, 2009). Unfortunately, though, this response threatens to generate trouble in the opposite direction. For if the relevant causal roles are specified too coarsely, critics foresee a rampant expansion of mind into the world. If I am often sat at my computer, do the contents of its hard drive count as my beliefs? What about the books on my office bookshelf or in the university library? Or the entire contents of the internet? Proponents of ExM would therefore seem to face a serious challenge when it comes to demarcating the boundaries of the mind. How can they specify the causal role for belief so as to allow for some cases of extended beliefs (like Otto) while avoiding a dramatic expansion of mind into the world? (Crane, 2016; Farkas, 2012; Sprevak, 2009).

Fictionalism offers a way to resolve this dilemma, since it rejects the idea that beliefs are individuated by their causal role. At the

198

heart of our talk about standing beliefs, the fictionalist insists, is not a definition of a causal role, but a metaphor – that of memory as a trusty, inner notebook. This metaphor can be apt and useful when applied to ordinary, biological memory, even if our biological memory exhibits effects that are absent from notebooks, like negative transfer. After all, this is a characteristic feature of metaphors: we use metaphors to relate a primary domain (e.g. memory, Italy) with a secondary domain (notebooks, boots), not to claim that they are exactly alike. The same lesson applies when we consider extended beliefs. The metaphor of memory as a notebook can be useful for understanding our interactions with external devices even when they function rather differently to written entries in notebooks, like pictures, maps or even smartphones. At this point, of course, critics will again see the risk of a dramatic expansion of mind into the world. How are we to draw a line around genuine cases of extended belief?

The correct response to this worry, I think, is to insist that there is no sharp line to be drawn. The aptness of any metaphor is a matter of degree. Consider the metaphor we use when we describe clouds as angry (Walton, 1993). There are cases in which everyone will agree that the metaphor is apt (say, standing in the midst of a force 10 gale) or that it isn't (gazing up at thin wisps of white cloud on a sunny day). In other cases, competent speakers may differ. The same is true of the metaphors that we use to describe the mind. In some cases, all will agree over whether these metaphors are apt or not. In other cases, there may be considerable disagreement. Indeed, even an individual speaker may feel uncertain if asked whether the metaphor is appropriate. Consider the numbers stored on a mobile phone. Mine includes the number of Exeter Library (01392 407027), but I can't recall this without using my phone. Do I believe that Exeter Library's phone number is 01392 407027? Well, yes and no. In many respects, the familiar metaphor of memory as a notebook captures my behaviour fairly well: if I need to call the library, I will call 01392 407027; if you ask me the area code, then I'll reply '01392'; and so on. In other respects, the metaphor is less apt: it'll probably take me a while to tell you the area code, as I'll need to look it up in my phone; at times when my battery is dead, I won't be able to tell you; and so on.

The key point is that, for the fictionalist, the boundaries of the mental are somewhat blurred: there are borderline cases in which there are aspects of someone's behaviour that our metaphors capture well and other aspects that it does not. Interestingly, Clark and Chalmers themselves acknowledge that there may not be 'categorical answers' to give regarding all putative cases of extended

belief ([1998](), p. 17). Of course, this will seem unsatisfying to those who want to draw a clear boundary between the mental and non-mental (cf. Sprevak, [2009](), pp. 522–23). And yet, I believe, this feature of fictionalism fits well with our ordinary talk about the mind, which is similarly ambiguous and conflicted. Consider the library phone number again. Suppose I'm late to give a talk at the library. In this context, the events manager might say, 'I wonder why he hasn't called us – he knows the number'. On the other hand, when I arrive I might offer an excuse: 'I'm sorry I couldn't call you – my phone was dead and I didn't know the number'. In these contexts, I submit, neither of these ways of speaking strikes us as especially odd or out of place. Both simply stress different aspects of the metaphor for different purposes.

Once again, notice that this is not a phenomenon unique to extended mental states. Even in cases where we do not rely on external devices, our talk about the mind can be equivocal in much the same way. Suppose that I didn't have Exeter Library's number stored in my phone, but still exhibited a similar pattern of behaviour: sometimes, with some effort and delay, I'm able to tell you the number; at other times, if I'm feeling tired or distracted, I struggle to recall it (cf. Farkas, [2012](); Sprevak, [2009]()). In this case too, I suggest, we would also hesitate and feel conflicted when asked if I knew the phone number (e.g. 'Come on, you know this!' versus 'Come on, you really ought to know this by now!'). Arguably, a similar ambiguity arises when we are asked about the beliefs of animals or pre-linguistic infants (e.g. Dennett, [2013](), pp. 96–97). Does Rover believe he lives at 99 Canine Avenue? Well, yes and no. Treating Rover as if he had a trusty inner notebook with his address written down will help to make sense of some aspects of his behaviour (e.g. that he heads back to 99 Canine Avenue if he gets off his lead) but not others (e.g. if you ask him his address, he's not especially forthcoming).

Conclusion

The extended mind thesis captures an extremely important insight into the role of external, material devices in transforming our minds. And yet, by retaining a traditional, representationalist view of mental states, it also faces serious criticisms and risks overlooking the way in which material culture shapes our vision of the mind as an inner realm. My aim in this paper has been to show that, by turning from representationalism to fictionalism, we can reach a new

200

understanding of the relationship between mind and materials. Of course, I have not offered a decisive argument in favour of fictionalism or against representationalism. There is much more that needs to be said to assess the relative merits of these approaches. But I hope at least to have shown that fictionalism offers a promising way to accommodate the important insights behind the extended mind thesis, while avoiding its more problematic consequences.[1]

References

Fred Adams & Ken Aizawa, 'The bounds of cognition', *Philosophical Psychology*, 14 (2001), 43–64.

Fred Adams & Ken Aizawa, *The Bounds of Cognition* (Oxford: Blackwell, 2008).

David Chalmers & Andy Clark, 'The Extended Mind', *Analysis*, 58 (1998), 7–19.

Andy Clark, *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing* (Cambridge, MA: MIT Press, 1989).

Andy Clark, *Associative Engines: Connectionism, Concepts, and Representational Change* (Cambridge, MA: MIT Press, 1993).

Andy Clark, *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* (Oxford: Oxford University Press, 2008).

Andy Clark, '*Memento's* Revenge: The Extended Mind, Extended', in Richard Menary (ed.), *The Extended Mind.* (Cambridge, MA: MIT Press, 2010a).

Andy Clark, 'Coupling, Constitution, and the Cognitive Kind: A Reply to Adams and Aizawa', in Richard Menary (ed.), *The Extended Mind.* (Cambridge, MA: MIT Press, 2010b).

Giovanna Colombetti & Tom Roberts. 'Extending the extended mind: the case for extended affectivity', *Philosophical Studies*, 172 (2015), 1243–63.

Tim Crane, *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation.* 3rd edition. (Abingdon, Oxfordshire: Routledge, 2016).

Tamás Demeter, (ed.). 'Mental Fictionalism' [Special Issue]. *The Monist*, 96(4) (2013).

Daniel Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press, 1987).

Daniel Dennett, 'Real Patterns', *The Journal of Philosophy* 88(1) (1991), 27–51.

Daniel Dennett, *Kinds of Minds: Toward an Understanding Of Consciousness* (New York: Basic Books, 1996).

Daniel Dennett, *Intuition Pumps and Other Tools for Thinking* (London: Penguin, 2013).

Ezequiel Di Paolo, 'Extended life', *Topoi*, 28(1) (2009), 9–21.

Douwe Draaisma, *Metaphors of Memory: A History of Ideas about the Mind* (Cambridge: Cambridge University Press, 2000).

Katalin Farkas, 'Two Versions of the Extended Mind Thesis', *Philosophia*, 40 (2012), 435–447.

Jerry Fodor, 'Where is my mind?', *London Review of Books*, 31(3) (2009), 13–15.

Ronald Giere, *Explaining Science: A Cognitive Approach* (Chicago: University of Chicago Press, 1988).

David Houghton, 'Mental Content and External Representations'. *The Philosophical Quarterly*, 47(187) (1997), 159–177.

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

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

Robert Rupert, 'Challenges to the Hypothesis of Extended Cognition', *The Journal of Philosophy*, 101(8) (2004), 389–428.

Gilbert Ryle, *The Concept of Mind*. (London: Hutchinson, 1949).

Eric Schwitzgebel, 'Belief'. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 edition). Retrieved from https://plato.stanford.edu/archives/sum2015/entries/belief/

Wilfrid Sellars, 'Empiricism and the Philosophy of Mind', in Herbert Feigl & Michael Scriven (eds.) *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*. Minnesota Studies in the Philosophy of Science, Vol. 1. (Minneapolis: University of Minnesota Press, 1956).

Mark Sprevak, 'Extended Cognition and Functionalism', *The Journal of Philosophy*, 106 (2009), 503–527.

Stephen Stich & Ted Warfield, (eds.) *Mental Representation: A Reader*. (Oxford: Blackwell, 1994).

Adam Toon, 'Fictionalism and the folk'. *The Monist*, 99 (2016), 280–295.

Adam Toon, 'Fictionalism and intentionality'. In Tamás Demeter, Ted Parent and Adam Toon (eds.) *Mental Fictionalism*. (Abingdon, Oxfordshire: Routledge, forthcoming).

Francisco Varela, Evan Thompson & Eleanor Rosch, *The embodied mind: cognitive science and human experience*. (Cambridge, MA: MIT Press, 1991).

Meg Wallace, 'Mental Fictionalism'. Unpublished manuscript, (2007).

Meg Wallace, 'Saving Mental Fictionalism from Cognitive Collapse', *Res Philosophica*, 93(2) (2016), 405–424.
Kendall Walton, 'Metaphor and Prop Oriented Make-Believe', *European Journal of Philosophy*, 1 (1993)., 39–57.
Daniel Weiskopf, 'Patrolling the Mind's Boundaries', *Erkenntnis*, 68(2) (2008), 265–276.
Michael Wheeler, 'In search of clarity about parity', *Philosophical Studies*, 152 (2011), 417–425.
Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1953. Translated by G. E. M. Anscombe).
Stephen Yablo, 'Does Ontology Rest on a Mistake?', *Proceedings of the Aristotelian Society, Supplementary Volumes*, 72 (1998), 229–261.

ADAM TOON (a.toon@exeter.ac.uk) *is Senior Lecturer in the Department of Sociology, Philosophy and Anthropology at the University of Exeter. He works in philosophy of science and philosophy of mind. He is the author of* Models as Make-Believe: Imagination, Fiction and Scientific Representation *(Palgrave Macmillan, 2012) and is currently writing a book on mental fictionalism.*

203