

*Draft of chapter that appears in Z. Radman (ed.), The Hand, an Organ of the Mind: What the Manual tells the Mental, Cambridge, Mass.: MIT Press, 2013, pp. 269-301.*

## **Is Cognition Embedded or Extended? The Case of Gestures**

**Michael Wheeler**

### **1. Locating Thought**

When we perform bodily gestures, are we ever literally thinking with our hands (arms, shoulders, etc.)? In the more precise, but correspondingly drier, technical language of contemporary philosophy of mind and cognition, essentially the same question might be asked as follows: are bodily gestures ever among the material vehicles that realize cognitive processes? More precisely still, is it ever true that a coupled system made up of neural activity and bodily gestures counts as realizing a process of thought, in such a way that the gross bodily movements concerned should be granted cognitive status along with, and in essentially the same sense as, the neural activity?

This question, however it is phrased, is, I think, acutely interesting in itself, but it enjoys the added value of bearing on the truth or otherwise of an increasingly prominent (although still very much minority) view in contemporary thinking about thinking, a view known as the *extended cognition (or extended mind) hypothesis* (henceforth ExC). ExC states that there are actual (in this world) cases of intelligent action in which thinking and thoughts (more precisely, the material vehicles that realize thinking and thoughts) are spatially distributed over brain, body and world, in such a way that the external (beyond-the-skull-and-skin) factors concerned are rightly accorded cognitive status. In this formulation, the term ‘cognitive status’ signals whatever status it is that we usually grant the brain within mainstream (internalistic, non-extended) cognitive theory, in relation to the production and explanation of psychological phenomena. According to ExC, then, under the right circumstances, the material vehicles that realize thinking and thoughts include not only neural states and processes, but also non-neural bodily structures and movements and, crucially (see next paragraph), certain environmentally located elements. Paradigmatic candidates for such environmental elements include a range of technological artefacts from notebooks and slide-rules to smartphones and other examples of contemporary mobile and pervasive computing. (The canonical statement of ExC is due to Clark and Chalmers (1998). Clark’s own recent presentation and defence can be found in (Clark 2008b). For a timely collection that

places the original Clark and Chalmers paper alongside a number of developments, criticisms and defences of the view, see (Menary 2010).)

For what it's worth, I am inclined to think that ExC is ultimately defensible (see, e.g., Wheeler 2010a) and so far more likely to be true than any knee-jerk first impressions of implausibility might suggest. However, even in the more revolutionary enclaves of contemporary cognitive theory, it is the fan of ExC who bears the burden of proof. To see why, we need to contrast ExC with a still-radical but ultimately less revisionist position on mind, which construes thinking not as extended, but as *embedded* (or *situated*). According to the embedded view, the distinctive adaptive richness and flexibility of intelligent behavior is regularly, and perhaps sometimes necessarily, causally dependent on (a) various non-neural bodily structures and/or movements, and/or on (b) the bodily exploitation of certain environmental props or scaffolds. Thus, to adapt an example from Clark (1997), consider an intelligent agent whose strategy for solving a jigsaw puzzle involves the systematic deployment of a range of bodily actions. For example, she might physically manipulate the pieces, either by rotating them to help visually pattern-match for possible fits or by testing them in the target position. Such problem-solving manipulations succeed by exploiting certain as-it-happens useful environmental factors, such as the geometric properties of the pieces. In the process, the external environment itself is effectively transformed into a readily available problem-solving resource, one whose elements restructure the piece-finding problem and thereby reduce the information processing load placed on the inner mechanisms in play.

As the jigsaw example shows, the embedded theorist seeks to register the important, and sometimes perhaps even necessary, contribution made by non-neural bodily and environmental factors to many cognitive outcomes. That said, the embedded position, as I shall understand it, is that the actual thinking in evidence in such cases remains *either* a purely neural phenomenon (one that is given a performance boost by its embodied context and its technological ecology) *or* a phenomenon that is realized in the brain and the non-neural body (but which is given a performance boost by its technological ecology). These two variations on the embedded view will be disentangled further in a moment. Right now, it is the distinction with ExC that concerns us. Here is the key point: if what we confronted in our jigsaw-solving scenario were a *full-blown* case of cognitive extension, as opposed to one of cognitive embeddedness, then not only the piece-manipulating bodily movements, but also the manipulated pieces themselves, would count, alongside and in the same fundamental sense as the active neural structures, as being among the material realizers of the thinking.

The case for embedded cognition has been made over and over again. (For what I now take to be my own book-length contribution to the cause, see (Wheeler 2005), although in hindsight I realise that, in line with the philosophical *Zeitgeist* of that era,

my treatment of the issues in that text fails to distinguish between the embedded view and ExC, in a sufficiently clear or consistent manner.) Given that ExC remains controversial in a way that the embedded view mostly doesn't, I think it is fair to say that it is the latter view that currently deserves to be treated as the default position in the debate. That's why the burden of proof rests with the advocate of ExC. Against this background, and taking it largely (although not entirely) for granted that the contribution of bodily gestures to thinking will surely indicate at least the embeddedness of cognition, my goal in what follows will be to determine whether or not a careful consideration of gestures and their role in thought mandates the more radical conclusion that cognition is extended. Putting gestures center stage in this way immediately highlights a subtle but important complication (or ambiguity) that was touched on in the last paragraph, and which is sometimes lost sight of in the literature. To bring this complication into clear view, consider the following, hypothetical turn of events: it turns out, when all the data and arguments are in, that although the properly cognitive part of our behavior-generating system includes non-neural as well as neural bodily factors, this process of cognitive spread stops at the skin. Under these circumstances, it seems to me that we would readily conceptualize cognition as a phenomenon that is embodied in previously unexpected ways, but not as a phenomenon that is extended in any truly interesting sense. That's why the formulation of ExC which I gave earlier is couched partly in terms of the cognitive status of *beyond-the-skull-and-skin* external factors.

If I am right about our inclinations here, then although the result that bodily gestures are among the realizing vehicles of cognition would certainly depose (what I shall call) the *conservative* embedded view, according to which extra-neural bodily factors and environmental structures are no more than powerful props for wholly neural cognitive activity, it would not *necessarily* be enough to establish ExC. For, without further argument, what that result would establish would be 'only' (what I shall call) the *radical* embedded view, according to which, although non-neural bodily factors sometimes count as being among the realizing material vehicles of cognition, the same is not true of any of the behavior-shaping factors located beyond the skin. What the fan of ExC needs to do, then, if she wants to build her case on gestures, is show that a robust implication of the very considerations that, when extracted from a careful analysis of gestures, bestow cognitive status on those gestures also bestow cognitive status on a range of extra-bodily environmental elements. Only then, it seems, would the cognitive status of bodily gestures provide proper support for a genuinely interesting and distinctive thesis of extended cognition. (Of course, what I am calling the radical embedded view might equally be dubbed minimal ExC, or alternatively be classified as a thesis of profoundly embodied or body-centric cognition; for more on body-centrism in relation to ExC, see (Clark 2008a, Wheeler 2011a). One's choice of terminology here depends on what aspect of the target position one wants to emphasize. In the end, what is important is that proper conceptual space is made for the position, not what we call it.)

With the substantive and somewhat lengthy scene-setting now over, it is time to get down to business. Here is where the rest of this chapter is going. I shall begin (in section 2) by pointing to certain empirical studies of bodily gestures which, on the face of things, provide a hefty experimental nudge in the direction of the claim that such gestures are indeed material realizers (more accurately, partial material realizers) of thought processes. I shall argue, however, that once we focus with due critical caution on the conceptual borders between the embedded view (in whatever form it manifests itself) and ExC, it becomes far less obvious that these empirical considerations alone mandate the eviction of the default (i.e., the embedded) view. To open the door to considerations that might conceivably complete the transition to a full-strength extended cognition framework here, I shall first say rather more about how my favored version of ExC works (section 3). Then, on the basis of that improved understanding of ExC, I shall proceed (in sections 4 and 5) to examine an argument due to Clark (2008b, this volume), which concludes explicitly that gestures are not merely props for subtly embedded but wholly neural cognitive activity, but instead are themselves literally part of cognitive processing, and then another argument, due to Gallagher (2005), which at least toys with that conclusion. In each case, I shall say why it is that even though the considerations on offer exhibit the right theoretical profile to make cognitive extension a genuine possibility, they ultimately fail even to secure the cognitive status of bodily gestures, which means that they don't tell in favor of ExC in contrast to the embedded view. Indeed, they leave us firmly in the grip of the conservative embedded view, rather than its radical cousin.

## **2. Experiments at the Edge of Embeddedness**

What sort of empirical data on bodily gestures might lead us in the direction of ExC? Here is a brief selection of some intriguing observations and experimental results. (Andy Clark's contribution to the present volume already contains a detailed summary of the first of my three examples, so I shall not repeat all its finer points here. Readers who hanker after more nuances are referred to Clark's chapter. Clark also discusses work that is closely related to my second example.)

Goldin-Meadow *et al.* (2001; see also Goldin-Meadow 2003) took two matched groups of children and asked them first to memorize a list, then to carry out a mathematical problem solving task, and then finally to recall the list. The children in one group were allowed to gesture freely during the mathematical problem solving phase. The children in the other group were instructed not to gesture. The outcome was that the children in the gesturing group did significantly better at recalling the list.

In another experiment, Broaders *et al.* (2007; see also Goldin-Meadow 2009) asked children to gesture while explaining their answers to novel mathematical problems. Children who could not solve the problems, and who did not give any spoken indication of having appropriate problem-solving strategies available to them, nevertheless indicated the possession of such strategies via their gestures. In this case, speech and gestures are vehicles for different meanings, with gestures carrying problem-solving content that, in an intuitive sense, is in advance of that carried by speech. This divergence seems to promote, or at least can be used to predict, later learning, since when the children in this gesturing group were later given an appropriate mathematics lesson, they were more likely to learn how to solve the problems in question than were children who, when explaining their preliminary, pre-teaching answers, had been asked not to gesture.

As a final example, consider Hutchins' (2010) analysis of a video record of an embodied communicative interaction between a company instructor pilot and his student. The context for this interaction is a training exercise in a flight simulator, in which the task is to land an airplane using the Instrument Landing System approach. After the student has made a poor landing, the instructor turns him and says "Want to do it again?," while performing a hand gesture on the words 'do it.' This gesture is interesting because it is not positioned over any of the controls or instruments on the simulated flight deck. It is a movement that starts at the current simulated position of the airplane on the runway and moves towards the back of the simulator, with the hand cupped and the palm facing backwards. Hutchins' analysis concludes that the gesture is in fact coupled to the simulated flight deck and thereby to the imaginary environment of an imaginary airplane in a partly simulated and partly imagined airspace. To explain: the referent of the anaphora 'it' in the spoken verb phrase "do it again?" is of course the flying of a simulated approach; but the referent of the accompanying gesture is a fictional event that would need to happen *before* another landing could be attempted, namely taxiing back up to the approach end of the runway in order to take off again. Indeed, the simulated airplane, unlike a real airplane, can be instantaneously repositioned at the start of an approach path, so there is no need for even a simulated taxiing and take-off phase. The direction of the gestural movement thus indicates the direction that a real airplane would need to travel in order to get back in the air, while the orientation of the hand corresponds to the orientation of the simulated approach path, in relation to the simulated flight deck of the imaginary airplane, as parked on the runway after the previous landing.

How are we to interpret these data? Goldin-Meadow (2003) suggests that the best explanation for the first result reported above is that the gestures performed during the mathematical problem solving task contribute actively to its execution. In this way, some of the cognitive load imposed by that task is transferred from the brain to the gestures, which leaves more of the available neural resources free for deployment on the accompanying memory assignment. In discussing the second

result, Goldin-Meadow writes that “[t]elling children to gesture . . . encourages them to convey previously unexpressed (and correct) ideas, which, in turn, makes them receptive to instruction that leads to learning” (Goldin-Meadow 2009, 108). Perhaps most striking of all, Hutchins (2010) argues that his instructor pilot’s gesture “indexes a conceptual construct that is a precondition for the concept indexed by talk” (96), and does so by “bring[ing] into being a conceptual space [a kind of scale model of the space of the approach path] in which [the gesture] acquires its own meaning” (98).

These various attempts to characterize the cognitive impact of gestures bring us tantalizingly close to the conclusion that at least some gestures are among the material realizers of thinking. In other words, we are currently located on the threshold of the radical embedded view and therefore firmly in the vicinity of ExC. Without further argument, however, it is not clear that the theoretical glosses in question actually force us to go beyond the conservative embedded position. For even if it is true that gestures contribute to thought in precisely the ways that Goldin-Meadow and Hutchins suggest they do, that still wouldn’t compel us (even though it might warmly invite us) to go beyond the view that gestures are ‘no more than’ powerful, non-cognitive causal influences on good old fashioned neurally instantiated thought. Here are some reflections that indicate why.

First, it seems clear that a processing or storage burden on my neural cognitive resources may be offloaded onto a non-neural resource without that latter resource *necessarily* acquiring cognitive status as a result. Consider, for example, the burden that used to be placed on my neural resources through the need to store important phone numbers in my brain. That load has more recently been relocated to my smartphone. And maybe that transferal has freed up some of my neural resources to do other things. But these facts *alone* surely don’t endow my smartphone with cognitive status, although they might be part of such a story. This deflationary (with respect to ExC) moral would seem to apply in just the same measure to the case of a processing or storage load being shifted onto bodily gestures.

Secondly, perhaps the solution-anticipating gestures performed by the children in Broaders *et al.*’s experiment are, in truth, bodily expressions of problem-solving strategies that have been unconsciously arrived at and stored in those children’s brains, but which the children concerned cannot yet express verbally. If this is indeed what is going on, then one might think that the gestures themselves are best understood as non-cognitive windows on the state of those children’s neural cognitive resources. As Clark (this volume, ??) puts this kind of point, perhaps “all the cognitive benefits of gesture are actually secured by a covert, fully neural, route.”

Finally, and here I confess that one has to squint with a certain amount of enthusiasm in order to ignore certain potentially disruptive details, compare (a) the

way in which the gesture performed by Hutchins' instructor pilot creates a conceptual space for a shared imaginary scenario, with (b) the way in which a software games designer or a contemporary film animator may use certain computational tools to create an imaginary world for public access. One intuitive gloss on (b), roughly analogous to the previously sketched offloading scenario, takes it that the computational tools available to the software world-builder constitute useful non-cognitive props and scaffolds that enable him to reduce the cognitive burden on his neural resources as he designs his imaginary space. A second gloss on (b), roughly analogous to the previously highlighted view that gestures are merely public expressions of fully formed neurally realized cognitive strategies, takes it that the world-builder, whether consciously or not, has the design of his imaginary space fully represented in his brain, in advance of that design merely being transferred to, by virtue of being implemented in, the appropriate software environment. Arguably, either gloss provides us with a deflationary yet satisfying interpretation of (b). And it is surely plausible that whatever goes for our software world-builder goes for Hutchins' gesturing instructor pilot.

From what we have seen so far, the interim conclusion is this: the (interesting and significant) contributions that gestures undoubtedly make to thought remain amenable to the conservative embedded picture according to which cognition is entirely neurally realized, even though it is intimately situated in the wider body and the world. In short, gestures help us to think, but (in the relevant sense) they are not themselves part of thinking. So much the worse, it seems, for the attempt to build a case for cognitive extension on the basis of gestures. Perhaps, however, in our rush to find empirical evidence for gesture-based cognitive extension, we are simply looking in the wrong place for the crucial considerations. In order to investigate that possibility, we need to bring ExC itself – or at least my favored version of it – into better focus. This demands that we temporarily turn our attention away from our main topic of gesture and towards more general questions of cognitive extension. The eventual pay-off from this detour, however, will be the uncovering of a theoretical structure, the so-called *mark of the cognitive*, that will enable us to see precisely what is needed by any gestures-based argument for ExC.

### 3. Riding the Waves

Broadly speaking, there are two prominent lines of thought in the ExC literature. Sutton (2010) dubs these the *first-wave* and the *second-wave* versions of the view. First-wave ExC is standardly characterized (by Sutton among others) as emphasizing and defending the kinds of arguments for cognitive extension that were to the fore in the original Clark and Chalmers (1998) paper. Almost all of the attention here is concentrated on the much discussed (and much misunderstood) *parity principle*. Clark's recent formulation of this principle is as follows: "[i]f, 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 (for that time) part of the cognitive process" (Clark 2008b, 77, drawing on Clark and Chalmers 1998, 8). So, the parity principle begins by asking us to consider an actual system that generates some psychologically interesting outcome (e.g., some example of intelligent action) and whose operation involves an important functional contribution from certain externally located physical elements. It then encourages us to imagine a hypothetical scenario in which exactly the same functional contribution, to an equivalent outcome, is made by certain internally located physical elements. Having taken this imaginative step, if we then judge that the internal realizing elements in the hypothetical case count as bona fide parts of a genuinely cognitive system, we are driven to conclude that the very same (i.e., cognitive) status should be granted to the external realizing elements in the actual, environment-involving case. After all, by hypothesis, nothing about the functional contribution of the target elements has changed. All that has been varied is the spatial location of those elements. And if someone were to claim that being shifted inside the head is alone sufficient to result in a transformation in status, from non-cognitive to cognitive, he would, it seems, be guilty of begging the question against ExC.

By contrast, so-called *second-wave* ExC rejects, or at least downplays, the parity principle, in favor of considerations of either *complementarity* (Sutton 2010; Kiverstein and Farina 2011) or, in a closely related vein, *cognitive integration* (Rowlands 1999; Menary 2007). As Sutton (2010, 194) argues, "in extended cognitive systems, external states and processes need not mimic or replicate the formats, dynamics, or functions of inner states and processes," so "different components of the overall (enduring or temporary) system can play quite different roles and have different properties while coupling in collective and complementary contributions to flexible thinking and acting." Adding a further dimension to, or perhaps making explicit an existing dimension of, complementarity, the integrationists emphasize the processes by which internal and external elements with different properties may be combined into a single, although essentially hybrid, cognitive whole. Thus they foreground factors such as the completion of cognitive tasks through the skilled manipulation of external elements, the transformation of our cognitive abilities through the learning of the manipulative skills just mentioned, and the application of norms of manipulation with a distinctively cognitive character (Menary 2007). The key, second-wave-defining commitment, however, is shared by complementarity theorists and their integrationist cousins. Put crudely, for the second-wavers, it's difference not sameness that matters.

Now, for my own part, I always did prefer the original punk bands to the so-called new wave bands that, in the context of the UK music scene anyway, came into being later (give me The Clash over The Cure any day), and things are not so different when it comes to ExC. So, although this is not the place to present a detailed critical



discussion of second-wave ExC, here is the kernel of an objection. According to the second-waver, it is precisely the differences between certain internal and certain external elements that explain how many cognitive tasks are performed. For example, following Bechtel (1994), one might explain how some examples of mathematical problem-solving are accomplished, by citing a complementary combination of internally located pattern-sensitive connectionist networks and externally located combinatorial symbol systems. One might even highlight the skilled embodied manipulation of the symbols in question, according to learned normative rules that, when mastered, transform what we can do. But none of this, as far as I can see (and here I am echoing the deflationary attitude to gestures that appeared earlier in our discussion), compels us to adopt ExC. At root, the problem for the second-wave theorist is this: once we remind ourselves that it is the embedded view which is the default position in the debate, and once we allow the inner elements in some proposed complementary combination of interest to have cognitive status (which, given my understanding of the notion of cognitive status, is not to say that the inner is necessarily always a self-standing cognitive system in its own right or that the cognitive is defined by whatever the inner does – see below), it seems that the second-wave emphasis on the existence of theoretically significant differences between the internal and the external elements in question creates a conceptual ravine between, on the one hand, the undoubtedly important phenomena of complementarity and integration and, on the other, cognitive extension. As Rowlands (2010, 90) puts it at the culmination of a similar line of reasoning, “given that there are significant differences between internal cognitive processes and external processes involved in cognition, why not simply suppose that the latter are part of the extraneous scaffolding in which the real, internal cognitive processes are *embedded*?” (In the interests of completeness, I should note that while Rowlands’ (1999) earlier work is standardly, and rightly, identified as one important source for integrationism, and thus as one wellspring of second-wave ExC, his more recent position (e.g. Rowlands 2010) has seen him argue that parity considerations (properly understood) and complementarity-integrationist thinking have equal weight in the justification for cognitive extension. Rowlands’ recent approach to parity is an issue to which I shall return briefly below.)

With at least the shape of an objection to second-wave ExC duly recorded, let’s refocus our attention on the first-wave version and the parity principle. Notice that, as stated above, the parity principle depends on the notion of *multiple realizability*, the idea that a single type of mental state or process may enjoy a range of different material instantiations. This dependence becomes visible (Wheeler 2011a) once one recognizes that the all-important judgment of parity is based on the claim that it is possible for the very same cognitive state or process to be available in two different generic formats – one non-extended and one extended. Thus, in principle at least, that state or process must be realizable in either a purely organic medium or in one that involves an integrated combination of organic and non-organic structures. In

other words, it must be multiply realizable. So, if we are to argue for cognitive extension by way of parity considerations, the idea that cognitive states and processes are multiply realizable must make sense. Now, one of the first things that undergraduate students taking philosophy of mind classes get taught is that the philosophical position known as *functionalism* provides a conceptual platform for securing multiple realizability. According to functionalism, as I shall understand it here, what matters when one is endeavoring to identify the specific contribution of a state or process *qua* cognitive contribution is not the material constitution of that state or process, but rather the functional role which it plays in generating cognitive phenomena, by intervening causally between systemic inputs, systemic outputs and other functionally identified, intrasystemic states and processes. This is the kind of functionalism that, I think, still deserves to be called the house philosophy of mind in cognitive science. For example, computational explanations of mental phenomena, as pursued in, say, most branches of cognitive psychology and artificial intelligence, are functionalist explanations, in this sense.

As an aside, notice that I am not proposing functionalism as a way of specifying the constitutive criteria that delineate the mental states that figure in our pre-scientific, folk (i.e., commonsense) psychology. Thus, for example, I am not advocating functionalism as a way of specifying what it is for a person to be in pain, as we might ordinarily think of that phenomenon. That ambitious brand of functionalism faces a range of well-documented objections that are widely thought to be fatal to the view. (For an introduction to the main lines of argument, see e.g. (Levin 2010).) It strikes me, however, that the functionalism that I have advocated in relation to ExC – call it *cognitive-scientific functionalism* – is pretty much immune to the objections traditionally leveled at the more ambitious project. Proper analysis and argument would undoubtedly be needed to make this point stick, but good empirical evidence for such an immunity is provided by the fact that the functionalist (in my sense) brands of artificial intelligence and cognitive psychology have not ground to a halt in response to the (alleged) failure of the folk-psychology-oriented version of the position.

Back to the main plot: Because a function is something that enjoys a particular kind of independence from its implementing material substrate, a function must, in principle, be multiply realizable, even if, in this world, only one kind of material realization happens to exist for that function. What this tells us is that functionalism is sufficient for some kind of multiple realizability. But is it *the right sort* of multiple realizability to support ExC? Historically, of course, the guiding cognitive-scientific assumption has been that the economy of functionally identified material states and processes that causally explain psychological phenomena will be realized by the nervous system or, in hypothetical cases of minded robots or aliens, whatever the differently realized counterpart of the nervous system inside the bodily boundaries of those cognitive agents turns out to be. In other words, functionalism has

standardly been used to secure (what we might call) *narrow* (within-the-skin) multiple realizability. In truth, however, there isn't anything in the letter of functionalism as a generic philosophical framework that mandates this exclusive focus on the inner (Wheeler 2010a, b). After all, what the functionalist schema demands of us is only that we specify the causal relations that exist between some target element and a certain set of systemic inputs, systemic outputs and other functionally identified, intrasystemic elements. There is no essential requirement that the outer boundary of the system of interest must fall at the organic sensory-motor interface. In other words, in principle at least, functionalism straightforwardly allows for the existence of material cognitive systems whose borders are located at least partly outside the skin. To put a different slant on this point, functionalism straightforwardly allows for (what we might now call) *wide* (both within-the-skin and beyond-the-skin) multiple realizability. And that is precisely what the first-wave ExC theorist needs in order to run the argument from parity. So, although I am happy to adopt Clark's term *extended functionalism* for the version of ExC that I favor (Clark 2008a, 2008b; see also Wheeler 2010a, 2010b, and 2011a), it is instructive to note that the term 'extended' in extended functionalism refers to the spatial limits of the material cognitive system of interest and not, as might perhaps be thought, to the theory of functionalism. In other words, the possibility of cognitive extension is not something that requires any theoretical enhancement of (any theoretical extension to) conventional cognitive-scientific functionalism. Rather, the possibility that ExC is true is a straightforward consequence of such functionalism. Put another way, if conventional cognitive-scientific functionalism is true, then ExC is not in any way conceptually confused, even though it may be empirically false. (For an argument which concludes that extended functionalism leads to a deadlock between ExC and its critics, see Rowlands 2010; for an argument which concludes that extended functionalism is committed to an excessively liberal notion of cognitive extension, see Sprevak 2009; for replies to both these arguments, see Wheeler 2010a.)

So far we have uncovered two conceptual components of (what I take to be) the most plausible form of ExC. But neither the parity principle nor functionalism, nor even the two of them combined, can carry the case for cognitive extension. What is needed, additionally, is an account of precisely which functional contributions count as cognitive contributions and which don't. After all, as the critics of ExC have often observed, and as the officially pro-ExC complementarity arguments canvassed above point out, there will undoubtedly be *some* functional differences between extended cognitive systems (if such things exist) and purely inner cognitive systems. For instance, our purely inner organic memory systems ordinarily exhibit primacy and recency effects that extended memory systems plausibly wouldn't, or at least needn't (Adams and Aizawa 2008; see Rupert 2009 for similar observations, and Wheeler 2010a, 2010b for discussion). So, faced with the task of deciding some putative case of parity, we will need to decide which, if any, of those functional differences matter. I see no secure way of doing this, except by providing what

Adams and Aizawa (2008) have dubbed a *mark of the cognitive*, a scientifically informed account of what it is to be a proper part of a cognitive system that, so as not to beg any crucial questions, is fundamentally independent of where any candidate element happens to be spatially located (Wheeler 2010a, 2010b, 2011a, 2011b). This way of explicating the basic idea of a mark of the cognitive specifies certain general conditions of adequacy that any particular suggestion for such a mark would need to meet. Of course, once a candidate mark of the cognitive has been placed on the table (i.e., once we have an account of what it is to be cognitive that meets the proposed adequacy conditions), further philosophical and empirical leg-work will be required to find out (i) whether that account is independently plausible, and (ii) just where cognition (so conceived) falls – in the brain, in the brain and the non-neural body, or, as ExC predicts will sometimes be the case, in a system that extends across brain, body and world.

In contrast to the foregoing picture, Clark has argued that the fan of ExC should shun the idea of a mark of the cognitive (as I have characterized it), in favor of “our rough sense of what we might intuitively judge to belong to the domain of cognition” (Clark 2008b, 114). Since this disagreement will figure, albeit in a minor way, later in our discussion, it is worth just pausing to introduce its main currents. According to Clark, judgments about whether or not some distributed system counts as an extended cognitive system should not be constrained by any scientific account of cognition, since such accounts are standardly “in the grip of a form of theoretically loaded neurocentrism” (Clark 2008b, 105). Rather, those judgments should be constrained by our everyday, essentially prescientific sense what counts as cognitive, since the “folk [i.e., commonsense] grip on mind and mental states . . . is surprisingly liberal when it comes to just about everything concerning machinery, location, and architecture” (Clark 2008b, 106). As a strategy for identifying cognitive structures and systems that doesn’t immediately beg the question against ExC, Clark’s appeal to folk intuitions strikes me as misguided. Indeed, as far as I can see, our ordinary folk practices of mental state attribution strongly presume the within-the-skin internality of our cognitive machinery (see Wheeler 2011b for the detailed argument; and see Clark 2011 for a response). Moreover, there is every reason to believe that it is possible to extract, from the functionalist varieties of artificial intelligence and cognitive psychology, certain candidate marks of the cognitive that, in their fundamental theoretical commitments, if not in the way they have actually been applied in practice, have not been seduced by any theoretically loaded neurocentrism (see, e.g., Wheeler’s 2011a interpretation of Bechtel’s aforementioned hybrid models of mathematical competence, as realizing extended physical symbol systems). This is, of course, precisely as one would expect, if my foregoing reflections on cognitive-scientific functionalism are on the right track.

At this point in the proceedings, something interesting happens. We are currently courting one of those Wittgensteinian ladder-discarding moments, in which we are

invited to throw away one of the very theoretical supports that got us to where we are. To explain: the role that is now being played, in our explication of ExC, by the idea of a mark of the cognitive threatens to remove the need for any appeal to parity. After all, if a candidate extended solution satisfies some agreed mark of the cognitive, it looks as if the argument for ExC is already complete and parity considerations are made redundant, meaning that the parity principle itself may be jettisoned. (Thanks to Peter Sullivan for pressing me on this point.) One response here would be to agree that, strictly speaking, the parity principle is not required as the engine room of first-wave ExC, but to point out that nevertheless it may continue to be a useful heuristic mechanism that helps to ensure equal treatment for different spatially located systems judged against an unbiased and theoretically motivated standard of what counts as cognitive. In short, it is a helpful bulwark against what Clark (2008b, 77) calls “biochauvinistic prejudice.” Another response would be to point out that a perfectly reasonable notion of parity may continue to figure in the method by which we appeal to our mark of the cognitive. Thus, some purely inner solution to a cognitive problem and some alternative, extended solution to the same problem may be judged to enjoy *parity with respect to a particular mark of the cognitive* (Wheeler 2010a, 2011a, 2011b; cf. Rowlands 2010, 90).

With our updated understanding of parity in hand, the first-wave ExC theorist is well-positioned to accommodate, within her theory, the kinds of differences in the formats, dynamics and functions of purely inner and extended systems that so impress both the second-wavers and some of ExC’s critics. The only differences between internal and external elements that, in the light of our revised parity considerations, would count against ExC would be those that, according to some particular mark of the cognitive in force, are stationed at the conceptual boundary between the cognitive and the non-cognitive (Wheeler 2010a, 2010b; cf. again Rowlands 2010, 90, although it should be noted that Rowlands doesn’t frame his project as a defence of first-wave ExC, but as the development of a view that he calls the ‘amalgamated mind’ which is designed to subsume embodied and extended approaches).

On the strength of the foregoing reflections, the notion of a mark of the cognitive can now be added alongside functionalism and the (properly understood) parity principle, in order to give us the conceptual profile of (the most plausible form of) ExC. Within this profile, a certain theoretical priority needs to be granted to the notion of a mark of the cognitive (or perhaps to the notion of a structured set of marks of the cognitive – one should admit that things are likely to get complicated!). ExC might well be able to survive without the parity principle (hence our ladder-discarding moment) and without functionalism (which is certainly sufficient but arguably not necessary for multiple realizability). But it is less clear that ExC could survive without the notion of a mark of the cognitive, given the way in which, without that notion in play, deflationary conservatism always favors an embedded

view. With that in mind, we can now be quite specific in stating what, in the overall context of this chapter, is the key issue. Any case for ExC based on gestures needs to provide a mark of the cognitive which has the consequence that bodily gestures are rightly counted among the material vehicles that realize cognitive states and processes. Moreover, to move us beyond the radical embedded view to full-blown cognitive extension (see above), that mark of the cognitive must plausibly allow various environmental elements to be accredited with cognitive status. In the next two sections, I shall consider two accounts from the recent embedded-extended literature which, although they do not frame things in quite the way I just have, may ultimately be interpreted as offering us just such a mark of the cognitive. In each case, however, I shall argue that even though the account in question is in the right theoretical ballpark for ExC – in the sense that if it did establish the cognitive status of gestures, it would generalize so as to secure, in addition, the cognitive status of various environmentally located elements – it fails to reach first base, because it does not establish the cognitive status of gestures.

#### 4. Cognitive Self-Stimulation

Clark (2008b, this volume) argues that the best way to understand the contribution of gestures to thought is to depict such activity as a form of cognitively potent *self-stimulation*. Here is the idea. In gesturing activity, neural systems are (partly) causally responsible for producing certain special-purpose bodily movements which are then recycled as inputs to those and/or other neural systems. This feedback process sustains sophisticated brain-body loops of exploitation, co-ordination and mutual entrainment, with various problem-solving benefits. In short, gestures are self-generated inputs to neural processing that (to borrow one of Clark's own analogies, more on which below) turbo-charge thought. In the interests of completeness, it is worth noting that Clark usefully distinguishes between three different species of cognitive self-stimulation: *fully anarchic*, *semi-anarchic* and *centrally controlled* (see Clark 2008b, 131-135). Cognitive self-stimulation is fully anarchic when it is realized by "a vast parallel coalition of more or less influential forces" whose unfolding is "largely self-organizing" (Clark 2008b, 131). It is semi-anarchic when it is realized by a system of loosely coupled autonomous mechanisms that, through purely local control protocols, are capable of exploring divergent trajectories without destructively interfering with each other. And it is centrally controlled when the self-stimulating activity of the system is orchestrated by some privileged executive system, canonically located in the brain. Clark argues that gestural cognitive self-stimulation could not be fully anarchic, because the gestural and verbal reasoning systems often require the 'protection' afforded by loose coupling, in order to explore different spaces of thought. (We saw experimental evidence of this effect earlier, in the experiment conducted by Broaders *et al.* (2007), in which children gestured while explaining their answers to novel mathematical problems.) Having ruled out the anarchic option, Clark allows that cognitive self-

stimulation may be either semi-anarchic or centrally controlled, with neither of these possibilities having any detrimental implications for the claim that gestures are among the material vehicles of cognition. In particular, he writes that it is “open to even the staunchest fan of central control to endorse [ExC]” (Clark 2008b, 244, note 27). Since Clark is, at root, neutral between the semi-anarchic and centrally controlled options, I shall not distinguish between them here and so will refer henceforth simply to the process of cognitive self-stimulation.

Clark’s compelling image of self-generated bodily movements acting as components in cognitively self-stimulating loops nicely captures what is distinctive about the kinds of gestures that, as we saw earlier, may plausibly enable the reallocation of neural resources, by soaking up some of the overall processing load, or may encode verbally inexpressible problem-solving strategies in ways that promote future learning. Less obviously, perhaps, the notion of cognitive self-stimulation also helps us to understand the case of Hutchins’ gesturing instructor pilot. Recall that, in this example, while the instructor’s speech serves to index a future simulated approach flight (that’s the reference of the word ‘it’ in the sentence “Want to do it again?”), the gesture which accompanies those words indexes the wholly imaginary event of the fictional airplane taxiing back up the runway in order to take off again. This imaginary event constitutes a precondition for the future simulated approach. Of course, unlike our other two examples of gesturing, the instructor’s bodily motion has a communicative function. But that shouldn’t blind us to the fact that the shared conceptual space it establishes also creates a platform for some otherwise unavailable individual reasoning. Hutchins (2010, 97) touches on this very point when he notes that the “gesture depicts a fictional event that facilitates reasoning,” and that it does so by bringing online complex visualization skills that instrument-rated pilots possess, and which support thinking and planning not only in flight simulators, but also when, for example, the view out of a real airplane is obscured by clouds. The instructor’s gesture may thus be seen as reflexively engaging his own visualization skills, as well as communicatively engaging those of his student.

What seems clear, then, is that gestures may act as self-generated aids that enhance thought in subtle and powerful ways. The question for us, though, is this: does the fact that gestures realize the phenomenon of cognitive self-stimulation provide ExC-related support for the claim that gestures are among the material vehicles that instantiate psychological processes? There seems little doubt that Clark thinks the answer to this question is “yes.” Indeed, he introduces his treatment of gesture as a “worked out example of extended cognizing in action” (Clark 2008b, 123). For Clark, then, cognitive self-stimulation is supposed to be the route by which gestures attain cognitive status. Unfortunately, once we look closely at the details of Clark’s argument, it is far from obvious that he delivers on this promise.

First, let's deal with what is ultimately a misguided objection to Clark's position. It wouldn't do for a critic here to insist that, because a self-generated input that figures in a self-stimulating loop remains an input, it cannot count as cognitive. Without further argument, this would only beg the question against Clark, who maintains that, *within the kind of self-stimulating loops at issue*, the fact that a self-generated input has the status of an input to a self-standing cognitive system (the neural system) is not a barrier to that input enjoying cognitive status. As Clark himself puts it, in a passage quoted in full below, "[i]n such cases [i.e., of cognitive self-stimulation], any intuitive ban on counting inputs as parts of [cognitive] mechanisms seems wrong" (Clark 2008b, 131). This point is reinforced by the analogy that Clark draws with a turbo-driven car engine, in which the exhaust flow is both an output and a self-generated input. There "can be little doubt that the whole turbocharging cycle [including the exhaust flow] should count as part of the automobile's own overall power-generating mechanism!" (Clark 2008b, 131). Similarly, we are invited to think, there should be little doubt that the whole cognitive self-stimulation loop, including the self-generated input (e.g. a pattern of bodily gestures), counts as part of the thinker's own cognitive mechanism. Interpreted carefully, as the specific claim that the self-generated input's standing as an input is, in itself, no impediment to it being awarded cognitive status, this point seems to me to be well-taken. The problem with Clark's treatment of cognitive self-stimulation must reside elsewhere.

To see where the problem lies, we need to pay detailed attention to precisely what Clark says about the character and the theoretical consequences of cognitive self-stimulation. Clark's cornerstone claim is that the "key distinction between "merely impacting" some inner cognitive process and forming a proper part of an *extended* cognitive process looks much less clear . . . in cases involving the systematic effects of *self-generated* external structure on thought and reason" (Clark 2008b, 126). To keep things in line here, we need to situate this claim in relation to the three-way distinction that I have been adopting, between the conservative embedded view, the radical embedded view, and ExC. The first thing to note is that, in the target passage from Clark, the term 'external structure' refers to extra-*neural* factors, so cognitive extension will be automatically secured if extra-neural factors count among the material realizing vehicles of cognition. One might think that this simply elides the distinction on which I have been insisting, between radical embeddedness and cognitive extension. However, because Clark is a good functionalist, he takes cognitive self-stimulation to be a process with respect to which it is fundamentally *irrelevant* whether the self-generated structure is a non-neural bodily factor or an environmental element. What this means is that, *if* the considerations he introduces support the radical embedded view, they will thereby support ExC, which is, of course, just another way of saying that Clark's appeal to cognitive self-stimulation is in the right theoretical ballpark to underpin ExC. Because Clark's position has this structure, it will be safe for us to explicate and interrogate that position largely in terms of a two-way distinction between the conservative embedded view and ExC.



For the rest of this section, then, I shall refer to the radical embedded view only when it is helpful to do so.

Using the conceptual machinery that we have just set out, we can see how Clark's cornerstone claim applies to bodily gestures in relation to thought. Clark's suggestion is that once the phenomenon of cognitive self-stimulation is brought into correct theoretical view, the idea that there is a clear boundary between the conservative embedded account of gestures (as non-cognitive factors that have an impact on neurally realized thought) and the extended account of gestures (as material vehicles of cognition) becomes far less compelling. It is at this point that the problem with Clark's position comes into focus. Strictly speaking, what Clark argues is not that, where cognitive self-stimulation is in evidence, we have a case of cognitive extension, but instead that, where cognitive self-stimulation is in evidence, *there is no clear distinction between the conservative version of cognitive embeddedness and cognitive extension*. Now, one might reasonably wonder why this constitutes an argument for the conclusion that ExC is to be preferred over the conservative embedded view, rather than an argument for the conclusion that there are circumstances in which whether one describes an action as a case of conservative cognitive embeddedness or as a case of cognitive extension is ultimately a matter of intellectual temperament rather than metaphysical correctness. Of course, Clark's own description of his treatment of gesture (see above) strongly indicates that this sort of even-handed outcome is not what he intends. But, at first sight anyway, that seems to be what he offers us.

Maybe we are missing something. Perhaps the right reconstruction of the target reasoning is this: in cases of cognitive self-stimulation, (i) the distinction between conservative embeddedness and cognitive extension is eroded in such a way that whatever evidence there is that tells in favor of the conservative embedded view, tells equally in favor of ExC, and (ii) under such circumstances, we are theoretically permitted to adopt ExC. Let's consider sub-claims (i) and (ii) in turn. One worry about sub-claim (i) is that it threatens to flout a point that has been made forcibly and repeatedly by Adams and Aizawa (e.g., 2008) in the ExC literature, namely that the bare causal dependence of thought and reason on external factors – even when that dependence is of a necessary kind or indicates that the systems in question are closely causally coupled – is simply not sufficient for genuine cognitive extension. What is needed, in addition, is a relation of *constitutive* dependence. That is, it must be that external factors don't merely exert a causal influence on, but rather partly constitute, the realizing base of cognition. Given this, the worry for Clark's sub-claim (i) is that the explicitly argued-for erosion of the distinction between conservative embeddedness and cognitive extension is being purchased using the dubious currency of a tacit and un-argued-for erosion of the more fundamental distinction between causal and constitutive dependence.

How might Clark respond? There is textual evidence which suggests that he believes the very same empirical evidence of cognitive self-stimulation that supposedly undermines the embeddedness-extension distinction also undermines the causal-constitutive distinction. Thus, in a passage from which we have already quoted, he writes:

*Sometimes, all coupling does is provide a channel allowing externally originating inputs to drive cognitive processing along. But in a wide range of the most interesting cases, there is a crucially important complication. These are the cases where we confront a recognizably cognitive process, running in some agent, that creates outputs (speech, gesture, expressive movements, written words) that, re-cycled as inputs, drive the cognitive process along. In such cases, any intuitive ban on counting inputs as parts of [cognitive] mechanisms seems wrong. (Clark 2008b, 131)*

As I interpret this passage, Clark's contention is that the empirical fact of cognitive self-stimulation establishes a context in which the causal-constitutive (or coupling-constitution) distinction has no force. If cognitive self-stimulation did indeed have this effect, my objection to sub-claim (i) would lose its force. But surely, as it stands, Clark's contention threatens to get things precisely back to front. For, given that we are in the business of pursuing a distinction-collapsing strategy, it is only by collapsing the causal-constitutive distinction that it becomes possible to take the empirically identified causal contribution of the self-generated inputs in question to be evidence of their constitutive status as material vehicles that partly realize thought. If my reasoning here is on track, then what Clark owes us, but doesn't give us, is an argument for collapsing the causal-constitutive distinction *that is independent of the empirical fact of cognitive self-stimulation*.

Even if we ignore the problems with sub-claim (i) of our reconstructed version of Clark's reasoning, sub-claim (ii) is questionable. In effect, this is the claim that, if the available evidence tells equally in favor of both the conservative embedded view and ExC, then we are theoretically permitted to adopt ExC. That would be true, of course, *if* the default position in the debate were ExC, but it is hard to see just what, in the present philosophical and scientific climate, might justify that assessment of the relative standings of the two theses. I pointed out earlier that, relative to ExC, the case for the embedded view is already well-established in the recent literature. And even if Clark responded to this observation by insisting that any empirical support for the conservative embedded view ought to be inherited by ExC in cases of cognitive self-stimulation (which is a putative pay-off from sub-claim (i)), the fact remains that ExC undoubtedly demands the more significant revision to our scientific and philosophical approaches to mind. In virtue of this genuine asymmetry, a perfectly reasonable theoretical inertia currently favors counting the

less revisionist, embedded option as the default view, with the conservative embedded view emerging as even harder to shift than its radical embedded cousin. This is all bad news for sub-claim (ii).

If I am right, our reconstructed version of Clark's reasoning falls doubly short. So, how might the appeal to cognitive self-stimulation be reconceived, so as to avoid the difficulties just highlighted? The reasoning that we have recently rejected fails partly because it either ignores or engages unsuccessfully with the causal-constitutive distinction. What would it be, therefore, to pay proper heed to, and thus to engage successfully with, that distinction? It would be to provide a set of constitutive criteria for a behavior-shaping element to count as a genuine material realizer of cognition. Readers who have been paying attention will right now be crying out that we have met this idea already, in the guise of the mark of the cognitive. So, here is a suggestion that would revitalize the appeal to cognitive self-stimulation: we should adopt the view that being the kind of self-generated input that supports a process of cognitive self-stimulation is a mark of the cognitive. If this understanding of cognitive self-stimulation is warranted, then, given the experimental and observational evidence that we have already reviewed, to the effect that gestures are indeed best understood as self-generated inputs that support a process of cognitive self-stimulation, then gestures will count among the material realizers of cognition.

As we have noted, the idea of a mark of the cognitive is something from which Clark has sometimes displayed a noticeable tendency to distance himself. It is worth just pausing, then, to register the fact that Clark's notion of what it is to be a self-generated input within a cognitively self-stimulating loop already meets the structural adequacy criteria identified earlier for being a mark of the cognitive. Thus, consider the following features of cognitive self-stimulation, as Clark thinks of it. First, it provides us with a *scientifically informed* proposal for identifying the cognitive: Clark motivates the notion by citing the kinds of experimental data on gesture that have been discussed in this chapter. Secondly, it's impressively *neutral regarding where the material vehicles that realize thought might be located*: Clark argues that not only bodily gestures, written words and overt speech (see earlier quotation from Clark 2008b, 131), but also neurally realized inner speech (see Clark 2008b, 135), may be vehicles for such self-stimulation. Finally, it is deployed by Clark precisely to *distinguish the cognitive from the non-cognitive*. In my mind at least, that settles it: if the pay-off here were to be a persuasive gesture-based argument for ExC, Clark himself is well-positioned to, and would have much to gain from, endorsing the view that being a self-generated input that supports a process of cognitive self-stimulation is a mark of the cognitive.

Unfortunately, there is a serious obstacle standing in the way of the proposed strategy. It seems that, just because some target element is a self-generated input that figures in a cognitive self-stimulating loop, that fact alone isn't *sufficient* for the

element in question itself to count as cognitive, because it may very well make its turbo-charging contribution to thought while remaining non-cognitive in character. The problem, then, is not that the self-generated inputs that figure in self-stimulating loops *cannot* be cognitive, but rather that an element may be a self-generated input that figures in a self-stimulating loop and *still not be cognitive*. This is just another way of saying that, even though a careful consideration of cognitive self-stimulation may give us a feature that meets the structural adequacy conditions for being a mark of the cognitive, nevertheless once our attention shifts to the specific content of the proposed candidate feature, it turns out to be an insufficiently robust indicator of the cognitive to play the kind of constitutive role being asked of it.

To bring this point (which is in truth no more than a modulation of the deflationary attitude that we encountered earlier) into proper view, we need to return, yet again, to Clark's text. Given the way in which we are currently understanding the target argument for ExC (i.e., as turning on the claim that being the kind of self-generated input that supports a process of cognitive self-stimulation is a mark of the cognitive), Clark's discussion contains a stark moment of what might be interpreted as either ambiguity or under-determination. Recall once more his claim that, in cases of cognitive self-stimulation, "any intuitive ban on counting inputs as parts of [cognitive] mechanisms seems wrong." Now notice that the term 'cognitive' is not present in Clark's original text. I inserted it for expository reasons, because the result is a passage that more conspicuously expresses what Clark intends. But if we now look at things with a duly critical eye, with Clark's own distinction-collapsing strategy found wanting, and with the causal-constitutive distinction in force, it is desperately unclear that the right to add the term 'cognitive' here – a right that Clark ultimately needs – has been properly earned. After all, both sides of the current debate (i.e., the conservative embedded side and the extended side) cheerfully accept that self-generated inputs that support cognitive self-stimulating loops operate within well-defined *mechanisms* that turbo-charge thinking. However, this observation does not do enough to establish that every element within the mechanisms concerned – and, in particular, the self-generated inputs in question – count as *vehicles that realize* thought, as opposed to *causal influences on* thought. For the turbo-charging mechanism at issue may very well be a hybrid system of cognitive and non-cognitive elements that interact causally so as to enhance overall psychological performance. Put another way, the problem is that, although we have an argument for the conclusion that the self-generated inputs at issue are proper parts of certain loop-shaped *mechanisms* that turbo-charge thought, we do not yet have an argument for the conclusion that those loop-shaped mechanisms are, *in their entirety and in their own right, cognitive mechanisms*. Perhaps the properly cognitive mechanisms in play are sub-systems of larger, performance-enhancing loops, where the latter are not cognitive mechanisms in their own right, even though they contain cognitive mechanisms. Putting the point another way, it is not that the kind of thought-enhancing self-stimulating loops on which we have been concentrating can

never be cognitive mechanisms in their entirety and in own right, but rather that, if sometimes they are, it's not their character *as* thought-enhancing self-stimulating loops that makes them so.

This claim may be bolstered if we reflect on an analogous, non-cognitive case. Consider: computer technology is now routinely used in many sports training regimes to improve the performance of elite athletes. Sometimes, this process of skill-enhancement happens by way of self-stimulating loops. For example, Baca and Kornfeind (2006) have designed a system in which individual rowers who are training on indoor rowing machines are monitored for factors that affect their technique. Real-time data displayed on a screen in front of the rowers (regarding, e.g., ground reaction and pulling forces) enables them to improve their rowing movements. In effect, the rowers lay down structures (patterns of bodily movement) that are then recycled as inputs to support them in tuning their bodily capacities in ways that will eventually result in improved performance on the water. This is, I submit, a self-stimulating loop, but one that that enhances not reasoning or cognitive learning, but the acquisition and honing of bodily skills. Now notice that although the feedback systems in place here, and in particular the self-generated inputs that the rower exploits, are core aspects of the *mechanism* by which the rower's body is tuned for improved performance, there is presumably little temptation to categorize those inputs as *realizers* of the observed bodily adaptation, as opposed to elements that have a *critical causal impact* on that adaptation. And it is not at all obvious why things should carve up any differently when the focus of attention is a self-stimulating loop that enhances thought. This is a shortfall that leaves ample room for the conservative embedded theorist to claim (with the additional weight provided by the default status of her view) that where the self-generated inputs in which we are interested are extra-neural, as in the case of bodily gestures, those structures are located outside the boundaries of the properly cognitive mechanisms in play, mechanisms which themselves remain neurally located.

Of course, this is not necessarily the end of the matter. It is open to the fan of cognitive self-stimulation to reply that it is only certain instances of the phenomenon that should be expected to secure cognitive extension. The hunt would then be on for the extra factor or factors that are additionally required. If we then allow ourselves to plug in the proposal that bodily gestures exhibit the extra factor or factors in question, the gestures-based argument for ExC would be squarely back in the game. With this sort of strategy in mind, it might be tempting to appeal to the fact that cognitive self-stimulation mechanisms in human behavior will standardly be the products of design – either by natural selection, development or learning. Thus Clark (2008b, 130) offers us three intuition-pumping examples: (i) a situation in which the rhythm of the rain on my window happens, by chance, to improve my thinking, by positively affecting the pacing, sequencing and timing of my thoughts; (ii) a robot that has been designed to exploit the rhythm of the rain to improve the

pacing, sequencing and timing of its reasoning operations; and (iii), following Dennett (1991), a self-stimulating spitting robot that has evolved to spit stored water in a rhythmic fashion at a metal plate on its body, in order to achieve the same outcome as in (ii). In essence, Clark's analysis of these examples is that (i) is (at best) a case of conservative embeddedness, (ii) is usefully categorized as a case of extended cognition, although some hesitation is due, and (iii) is a clear case of extended cognition. The feature that makes example (ii) a (perhaps precarious) case of extended cognition is supposed to be the fact that the robot has been *designed* to exploit the external structures in question. That feature – the designed exploitation of external structures – is then presumably solidified and enhanced through the addition of the (also designed) self-stimulation mechanism described in (iii).

On the strength of Clark's brief discussion of the role that design may play in an argument for ExC, it seems reasonable to give serious consideration to the following thought: if the bare fact of cognitive self-stimulation isn't enough to guarantee ExC, perhaps the recognition that cognitive self-stimulation will standardly be a product of design (by evolution, development or learning) is. A moment's reflection, however, reveals a serious difficulty with this proposal, namely that *having been designed is not a robust mark of the cognitive*. For example, the heart has been designed by natural selection to pump blood around the body, but its designed character does not in any way make it a realizer of cognitive states or processes. So, merely pointing out that a material system or element which has not already attained cognitive status has been designed to perform its function will not result in the desired conceptual 'upgrade'. What is still missing is an account of why the designed function of interest counts as a cognitive function. The lesson for the proposed extended account of gestures is seemingly straightforward. Given that merely being causally coupled to an existing cognitive system such as the brain is not sufficient for cognitive status (that's one implication of the causal-constitution distinction), and given that (as I have argued) being a self-generated structure that supports a cognitive self-stimulation loop is not sufficient for cognitive status, simply adding in the thought that the loop and its components are designed features is of no relevant consequence. So, even if, as seems likely, bodily gestures are indeed designed structures of the appropriate kind, we still find ourselves some way short of ExC.

## 5. Prenoteic Constraints

Time for a change of tack. In this section, I shall examine certain central threads in a detailed treatment of bodily gestures due to Gallagher (2005). Gallagher, it must be admitted, does not quite come out and say that gestures are among the material realizing vehicles of cognition, but he does openly flirt with that thought, and it would not be unreasonable to interpret his arguments as warmly recommending an ExC-friendly outcome. Thus, in a passage also quoted in part and commented on by Clark (2008b, this volume), Gallagher writes as follows:

The question here is . . . about the cognitive effects gestures might have even if we have no conscious access to them. This is an extremely difficult question to answer if we think of cognition (thought) as a completely internal process that happens in a disembodied mind. It may be, however, that certain aspects of what we call the mind just are in fact nothing other than what we tend to call expression, that is, occurrent linguistic practices ('internal speech'), gesture, and expressive movement. (Gallagher 2005, 121, footnote 7)

Henceforth, I shall ignore Gallagher's hesitancy and assess the considerations he offers us as a gesture-driven case for ExC.

Gallagher provides (what he dubs) an *integrative* theory of gesture, one that understands gesture to be (i) *embodied*, in the sense of being facilitated and constrained by motor capacities, (ii) *communicative*, in the sense of being used in intersubjective co-ordination and communication, and (iii) *cognitive*, in the sense (or senses – see below) of “contributing to the accomplishment of thought” and “shaping the mind” (Gallagher 2005, 123). It is of course (iii) and its accompanying locutions that concern us. In accordance with our strategic orientation, I shall treat (iii) as an attempt to offer us a candidate mark of the cognitive. Let's begin, then, by considering the idea that gestures are genuine realizers of cognitive processes because they are elements that *shape the mind*.

Gallagher unpacks the relevant notion of 'shaping' by way of his concept of a *prenoetic* contribution to thought. A contribution to thought is prenoetic when it shapes (or structures) thought, but “does not normally enter into the phenomenal content of experience in an explicit way” (Gallagher 2005, 2). Here are two illustrative examples given by Gallagher (2005, 2-3). First, because of my nature as an embodied being, my perceptual access to the world necessarily takes place from a particular and limited spatial perspective. That perspective shapes what I experience, but it is not normally something of which I am explicitly conscious in my experience. Secondly, it is arguable that I always experience the world through an affective lens formed by my current mood. Roughly, if I'm depressed, the world strikes me as a somber place; if I'm euphoric, the same elements that might have had a gloomy hue strike me as no more than mildly troublesome irritants. My affective lens shapes what I experience, but it is not normally something of which I am explicitly conscious in my experience. For Gallagher, bodily gestures are further examples of such prenoetic factors. As he puts it, gestures “shape cognition in a prenoetic manner” (Gallagher 2005, 123).

To do justice to Gallagher's talk of 'shaping' here, and in the light of his examples of our limited spatial perspective and our affective lens, one might reasonably

conceptualize prenoetic contributions to thought as *constraints on* thought, or at least on thought as we know it. But once this interpretation of the prenoetic is in place, any claim that such factors, wherever they happen to be located, are material realizers of cognition, as opposed to important causal determinants of cognition, is less than compelling. Once again it is expositively useful to consider an analogous, non-cognitive example. It is widely recognized that the biological process of adaptation by natural selection operates against a backdrop of various constraints. For example, developmental constraints are imposed by the heavily conserved Hox genes which are active determinants in body segmentation and organ development in the anterior-posterior body pattern of many animals, including humans. In other words, basic bodily form – roughly, where the head goes, where the legs go, etc. – is a developmental constraint on adaptation. Practically speaking (i.e. outside of thought experiments), this sort of constraint cannot (or at least can very rarely) be overcome by selection, even if an obvious adaptive benefit would accrue. So one might picture the set of such constraints as taking a space of conceivable phenotypes (variation that we can imagine) and reducing it to a space of possible phenotypes ('possible' in the sense of variation that is actually available to selection). By channeling selection in this way, the constraints in question may be said to shape the adaptations that actually come about. Roughly, just as our embodiment constrains our spatial perspective and thereby the structure of our experience, our Hox genes may constrain, for example, the number of digits we can evolve on each hand and thereby the structure of our bodily adaptations. Crucially, once we have classified a biologically relevant factor as a constraint on some adapted trait of interest, and so have accepted that there is a fundamental sense in which that factor is simply not available for modification by selection, it would surely be theoretically uncomfortable to think of it as a proper part of, or as realizing, the adapted trait itself. Similarly, once we have classified a psychologically relevant factor as a prenoetic element with regard to some cognitive trait of interest, and if we have accepted that prenoetic elements shape cognition in the sense of being constraints on the structure of cognition, it would surely be equally theoretically uncomfortable to think of the prenoetic factor in question as a proper part of, or as realizing, the cognitive trait itself. The upshot, then, is that even if Gallagher is right that gestures prenoetically shape thought, that does not make gestures realizing material vehicles of cognition.

What about the variant understanding of gestures (as cognitive) that Gallagher offers us? This is the claim that gestures contribute to the *accomplishment* of thought. As Gallagher (2005, 121) makes clear, our source of illumination here is Merleau-Ponty's account of language as accomplishing thought, an account which turns on the somewhat nebulous notion of the *expression of thought as the completion of thought*. As Merleau-Ponty (1945/1962, 206) puts it:



If speech presupposed thought, if talking were primarily a matter of meeting the object through a cognitive intention or through a representation, we could not understand why thought tends towards expression as towards its completion, why the most familiar thing appears indeterminate as long as we have not recalled its name, why the thinking subject himself is in a kind of ignorance of his thoughts so long as he has not formulated them for himself, or even spoken and written them, as is shown by the example of so many writers who begin a book without knowing exactly what they are going to put into it.

Gallagher (2005, 121) builds on Merleau-Ponty's thinking about language to suggest that gesture, itself conceptualized as a mode of language, assists in the accomplishment of thought. In this Merleau-Pontian register, then, Gallagher's claim may be glossed in the following way: *gestures are expressions of thinking which complete that thinking*.

One can see how, in general terms, this picture might plausibly have the implication that gestures are material realizers of cognition. For, where this picture applies, an unexpressed thought is an incomplete thought, so the expression (e.g. speech or gesture) is plausibly a proper part of the thinking. This may well be what drives Gallagher's suggestion that "certain aspects of what we call the mind just are in fact nothing other than what we tend to call expression." What is still missing, though, is a proper account of what it is for an element to express thought in such a way that the process of expression accomplishes the completion of the thinking.

Further light may be shed on this matter if we adapt, for our own purposes, Krueger's (forthcoming) useful distinction between three different ways in which the notion of 'expressing thought' might be understood. As we shall see, Krueger's analysis makes good use of our old friend, the causal-constitutive distinction. The first option that Krueger identifies is that expressive behavior may be understood as expressing thought, because such behavior is the *causal output* of certain internal psychological phenomena. It should be clear that, as it stands, this understanding of what it is to express thought will not be adequate to ground a notion of cognitive accomplishment that succeeds in conferring cognitive status on the expressing elements. For the proposal gives us no reason to classify the relevant causal outputs of some inner psychological phenomenon as anything other than non-cognitive structures that are associated with that phenomenon. On the second option that Krueger identifies, the expressed psychological phenomenon is held to be *experientially co-present* with the expressive behavior associated with that phenomenon. The inspiration for this suggestion is the phenomenological point that when I visually perceive, say, a tomato, although my visual access to that entity is aspectual (there is an obvious sense in which, given my embodied spatial perspective, I have perceptual access only to certain portions of it), my experience is

of the tomato as an intact, solid, three-dimensional object. The tomato's hidden-from-view aspects are not known inferentially, but are experientially co-present with those aspects that are perceptually present. If we transpose this model to the present context, we get the following model: although what we perceive is the expressive behavior, such as the gesturing, we simultaneously experience the associated psychological phenomena. As Krueger himself notes, there is undoubtedly a *prima facie* disanalogy here, since whereas I could turn the tomato around to bring its currently-unperceived aspects into view, no similar action is available to me, within the co-presence model, in the case of the psychological phenomena that are associated with expressive behavior. What is interesting about this disanalogy in the present context is that it points us in the direction of a crucial observation that Krueger proceeds to make, namely that, on the co-presence model, when I observe another's gesturing I do not observe realizing material vehicles of her thinking, but rather movements that are associated with such vehicles. Once again, then, the proposed understanding of what it is to express thought will not be adequate to ground a notion of cognitive accomplishment that succeeds in conferring cognitive status on the expressing elements.

The third and final option that Krueger identifies takes expressive behavior to express thought in the sense that the behavior in question is *partly constitutive* of the psychological phenomenon of which it is an expression. Now, this is a notion of expressing thought that fits with Gallagher's suggestion that expressive movements help to accomplish thinking by expressing thought in the sense of completing the thinking in question. Of course, given that the constitutive dependence of thought on expressive bodily movements is precisely what is being advocated here, the proposal on the table will certainly support the conclusion that bodily gestures are among the material realizers of cognition. Moreover, there seems to be little doubt that environmentally located elements may partly realize expressive activity, so the account is plausibly ExC-compliant. Perhaps, then, the route to a gestures-based argument for ExC is to be found in Gallagher, via Merleau-Ponty and Krueger.

Regrettably, we are not quite home. By appropriating Krueger's analysis in the way that we have, we have certainly learned that if we want to maintain both the claim that gestures are expressions of thought and the claim that gestures partly realize thought, then we will need to understand the notion of expression in Krueger's constitutive sense. But while that is an important outcome, it is in truth no more than a reminder that a mark of the cognitive – a constitutive account of what makes a state, structure or process a cognitive one – is required. In the present context, such an account would specify the function or functions that an expressive behavior would need to perform, in order to count as completing, rather than just as influencing, thought. In fact, Krueger's own analysis implicitly reflects, although it does not explicitly make, this very point. For he proceeds to complete his account of expressive behavior as partly constitutive of thought by appealing to Clark's notion

of cognitive self-stimulation (Krueger forthcoming, 27). Unfortunately, as I argued earlier in this chapter, the fact that the relevant kinds of behavior exhibit the phenomenon of cognitive self-stimulation ultimately fails to plug the causal-constitutive gap. What all this indicates, I think, is that Gallagher's Merleau-Pontian suggestion that gestures help to accomplish thought – i.e., that gestures express thought so as to complete thought – is perhaps best seen as a place-holder for a mark of the cognitive, rather than as a mark of the cognitive itself. And that means that we still remain some way short of a gestures-based argument for ExC.

## **6. An Unfinished Business**

The conclusion of our investigation may be put as follows. If we ask ourselves the question 'Is cognition embedded or extended?', and we appeal only to the gesture-related arguments from the recent literature canvassed in this chapter, our answer ought to be 'embedded'. Moreover, it ought to be 'embedded' in the conservative sense of that term. But while this is in some ways a disappointing result for those of us who are fans of extended cognition, it's not all doom and gloom. Along the way we have learned some useful lessons. Most notably, we have learned that, once the extended cognition hypothesis is understood, as I think it should be, in terms of a tripartite profile involving the parity principle, extended functionalism, and the mark of the cognitive, certain gesture-related considerations, such as cognitive self-stimulation, prenoetic shaping, and expression as completion, are seemingly unable to carry the weight of argument that has sometimes been placed upon them. Of course, other considerations (other marks of the cognitive) may ultimately confer cognitive status on gestures, and yet other considerations (not applicable to gestures) may persuade us (or most of us anyway) that cognitive extension is indeed the way of things. Right now, however, it's an unfinished business – and that's the most interesting kind.

## **Acknowledgments**

Many thanks to Andy Clark, Shaun Gallagher, Zdravko Radman, and audiences in Dubrovnik, Bielefeld and Milan for important critical feedback on the ideas presented here. Some of the text in section 3 was adapted from passages that appear in (Wheeler 2011a, 2011c).

## **References**

- Adams, F. and K. Aizawa. 2008. *The Bounds of Cognition*. Malden, MA and Oxford: Blackwell.
- Baca, A. and P. Kornfeind. 2006. Rapid Feedback Systems for Elite Sports Training. *IEEE Pervasive Computing* 5 (4): 70-76.

- Bechtel, W. 1994. Natural Deduction in Connectionist Systems. *Synthese* 101: 433-463.
- Broaders, S., S. W. Cook, Z. Mitchell, and S. Goldin-Meadow. 2007. Making Children Gesture Brings out Implicit Knowledge and Leads to Learning" *Journal of Experimental Psychology: General* 136(4): 539–550.
- Clark, A. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA.: MIT Press.
- Clark, A. 2008a. Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind? *Philosophy and Phenomenological Research* 76 (1): 37-59.
- Clark, A. 2008b. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press.
- Clark, A. 2011. Finding the Mind. *Philosophical Studies*, symposium on Clark's *Supersizing the Mind*, 152(3): 447-461.
- Clark, A. and D. Chalmers, D. 1998. The Extended Mind. *Analysis* 58 (1): 7-19.
- Dennett, D. C. 1991. *Consciousness Explained*. Boston: Little, Brown.
- Gallagher, S. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.
- Goldin-Meadow, S., H. Nusbaum, S. Kelly, and S. Wagner. 2001 . Explaining Math: Gesturing Lightens the Load. *Psychological Science* 12: 516-522.
- Goldin-Meadow, S. 2003. *Hearing Gesture: How our Hands Help us Think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S. 2009. How Gesture Promotes Learning Throughout Childhood. *Child Development Perspectives* 3: 106–111.
- Hutchins, E. 2010. Imagining the Cognitive Life of Things. In *The Cognitive Life of Things: Recasting the Boundaries of the Mind*, L. Malafouris and C. Renfrew, eds., 91-101. Cambridge: McDonald Institute Monographs.
- Kiverstein, J., and M. Farina. 2011. Embraining Culture: Leaky Minds and Spongy Brains. *Teorema* 32 (2): 35-53.
- Krueger, J. (forthcoming). Seeing Mind in Action. *Phenomenology and the Cognitive Sciences*, special issue on empathy and intersubjectivity.

- Levin, J. 2010. Functionalism. *The Stanford Encyclopedia of Philosophy*, (Summer 2010 Edition), E. N. Zalta, ed., URL = <http://plato.stanford.edu/archives/sum2010/entries/functionalism/>. Last accessed October 10, 2011.
- Menary, R. 2007. *Cognitive Integration: Mind and Cognition Unbounded*. Basingstoke: Palgrave Macmillan.
- Menary, R., ed. 2010. *The Extended Mind*. Cambridge, MA: MIT Press.
- Merleau-Ponty, M. 1945/1962. *Phenomenology of Perception*. London and New York: Routledge.
- Rowlands, M. 1999. *The Body in Mind*. Cambridge: Cambridge University Press.
- Rowlands, M. 2010. *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge, MA: MIT Press.
- Rupert, R. 2009. *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.
- Sprevak, M. 2009. Extended Cognition and Functionalism. *The Journal of Philosophy* 106: 503–527.
- Sutton, J. 2010. Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process. In Menary (2010) 189-225.
- Wheeler, M. 2005. *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.
- Wheeler, M. 2010a. In Defense of Extended Functionalism. In Menary (2010), 245-270.
- Wheeler, M. 2010b. Minds, Things, and Materiality. In *The Cognitive Life of Things: Recasting the Boundaries of the Mind*, L. Malafouris and C. Renfrew, eds., 29-37. Cambridge: McDonald Institute Monographs. To be reprinted in *Action, Perception and the Brain: Adaptation and Cephalic Expression*, J. Schulkin, ed. Basingstoke: Palgrave Macmillan.
- Wheeler, M. 2011a. Embodied Cognition and the Extended Mind. In *The Continuum Companion to Philosophy of Mind*, J. Garvey, ed., London: Continuum, 220-238.

Wheeler, M. 2011b. In Search of Clarity about Parity. *Philosophical Studies*,  
Symposium on Andy Clark's *Supersizing the Mind* 152(3): 417-25.

Wheeler, M. 2011c. Thinking Beyond the Brain: Educating and Building, from the  
Standpoint of Extended Cognition. *Computational Culture* 1,  
<http://www.computationalculture.net/>