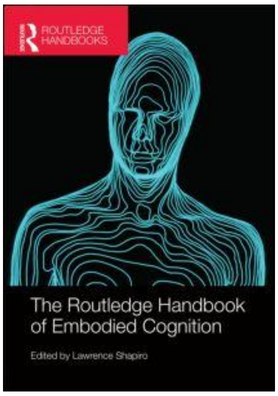

The Routledge Handbook of Embodied Cognition

Lawrence Shapiro

Revolution, Reform, or Business as Usual?

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9781315775845.ch35
Michael Wheeler
Published online on: 01 May 2014

35

REVOLUTION, REFORM, OR BUSINESS AS USUAL?

The future prospects for embodied cognition

Michael Wheeler

A pre-revolutionary situation?

When all the data and arguments are in, will the recent flurry of work in embodied cognition deliver a revolutionary paradigm shift in the sciences and philosophy of mind? Or will it be a case of business as usual in the mind-targeting laboratories and armchairs around the globe? Or is the most likely outcome a reformist tweak in which embodied cognition research is recognized as making genuine and important methodological or orienting contributions to cognitive science, while leaving the most fundamental conceptual foundations of the field intact – as Rupert nicely puts it in his sobering set of conclusions regarding the revolutionary implications of embodied approaches in general, "more of a nudging than a coup" (Rupert, 2009, p. 242)?

Reaching a judgment on this issue is trickier than it may at first appear, since it is not simply a matter of working out whether or not some homogeneous and well-delineated new research program in cognitive science has succeeded, or will succeed, in substantively reforming, or even deposing, the orthodox view in the field. (For present purposes, the orthodox view may be located by its adherence to the principle that intelligent thought and action are standardly to be explained in terms of the building and manipulation of inner representational states by inner computational processes.) Indeed, if we consider the assortment of thinkers who congregate under the banner of embodied cognition, and we reflect on the range of theoretical approaches on offer, it turns out that they constitute a far from univocal crowd. Because of this, my goal in this chapter is not to pronounce on the future prospects for embodied cognition research *in general* – that would require at least one book – but rather to evaluate the prospects for a limited number of specific but prominent views that march under that banner. In each case I shall endeavor to get clear about the fundamental theoretical commitments of the view in relation to a specific diagnostic issue, namely what its advocates mean by the presumably foundational notion of being "embodied."

The implementational body

According to the hypothesis of *extended cognition* (ExC), there are actual (in this world) cases of intelligent thought and action in which the thinking and thoughts (more precisely, the material vehicles that realize the 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.[1] To bring ExC into proper view, it is useful to compare it with an adjacent position in intellectual space, namely the hypothesis of *embedded cognition*. According to this second and arguably less radical position, intelligent thought and action are regularly, and perhaps sometimes necessarily, causally dependent on the bodily exploitation of certain external props or scaffolds.[2] So, consider the example of a mathematical calculation achieved, in part, through the bodily manipulation of pen and paper. For both the embedded view and ExC, what we have here is a brain-body-pen-and-paper system involving a beyond-the-skin element that, perhaps among other things, helps to transform a difficult cognitive problem into a set of simpler ones (e.g. by acting as storage for intermediate calculations). For the embedded theorist, however, even if it is true that the overall mathematical problem could not have been solved, at least by some particular mathematician, without the use of pen and paper, nevertheless the external resource in play retains the status of a non-cognitive aid to some internally located thinking system. It certainly does not qualify as a proper part of the cognitive architecture itself. In other words, the thinking in evidence remains a resolutely inner, paradigmatically neural, phenomenon, although one that has been given a performance boost by its local technological ecology. By contrast, for the advocate of ExC, the coupled system of pen-and-paper resource, appropriate bodily manipulations, and in-the-head processing *may* itself count as a cognitive architecture, even though it is a dynamically assembled (rather than hard-wired) and essentially temporary (rather than persisting) coalition of elements. In other words, each of the differently located components of this distributed (over brain, body and world) multifactor system enjoys cognitive status, where the term "cognitive status" should be understood as indicating whatever status it is that we ordinarily grant the brain in mainstream cognitive science. Another way to put this is to say that, according to the embedded view, the dependence of cognition on environmental factors is "merely" causal, whereas, according to ExC, that dependence is constitutive (Adams and Aizawa, 2008).

ExC naturally qualifies as a species of embodied cognition, because it takes non-neural bodily factors (e.g. manipulative movements of the body) to be parts of the realizing substrate of cognitive phenomena. Given our present remit, however, it is instructive to bring out the precise nature of the style of embodiment that is on offer. Here is an orienting thought. Surprisingly, perhaps, the possibility of extended cognition is a straightforward consequence of what still deserves to be labeled the house philosophy of mind in cognitive science, namely *functionalism*. The cognitive-scientific functionalist holds that what makes a state or process cognitive is not its material composition, but the functional role it plays in generating psychological phenomena, by intervening causally between systemic inputs, systemic outputs and other functionally identified, intrasystemic states and processes. Computational explanations, as pursued in, say, cognitive psychology and artificial intelligence (AI), are functionalist in this sense. Of course, historically, the assumption in cognitive science has been that the organized collection of states and processes that the functionalist takes to be the machinery of mind will be realized by the nervous system (or, in hypothetical cases of minded robots or aliens, whatever the counterpart of the nervous system inside the bodily boundaries of those cognitive agents turns out to be). In truth, however, there isn't anything in the letter of functionalism that mandates this internalism (Wheeler, 2010a, 2010b). After all, what functionalism demands is 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. Nothing here demands internalism, since the boundaries of the functionally identified system of interest – i.e. the cognitive system – may in principle fall beyond boundaries of the organic sensorimotor interface.

Clark (2008a, 2008b; followed by Wheeler 2010a, 2010b, 2011a) uses the term *extended functionalism* to describe the combination of functionalism and ExC. This moniker is useful, as long as one remembers that the qualification "extended" attaches to the nature of cognition and not to the thesis of functionalism. Nothing about that venerable philosophical thesis has changed, since the claim that cognition might be extended merely unpacks one implication of the functionalist picture that had been there all along. As one might put it, ExC, if true, is simply a footnote to functionalism. If we look at things from the other direction, however, the alliance with functionalism gives the ExC theorist something she needs – assuming, that is, that she wants to hold onto the presumably attractive thought that the very same type-identified cognitive process may, on some occasions, take place wholly inside the head, while on others it may take place in an extended brain-body-environment system. To explain: even if some mathematical calculations are simply too difficult for me to complete entirely in my brain, there are others for which it seems plausible to say that, although on Monday I may carry them out using pen and paper, on Thursday I may call only on my organic resources. Now, if we are to describe these alternative problem-solving routines as two realizations of the very same mathematical cognition, it must be possible for the very same psychological reasoning process to enjoy (at least) two different material instantiations. In other words, the target cognitive phenomenon must be *multiply realizable*. And while functionalism may not be necessary for multiple realizability, it is standardly thought to be sufficient, since a function is something that enjoys a particular kind of independence from its implementing material substrate. Indeed, it seems plausible that anything worthy of being a function must, in principle, be multiply realizable.[3]

Functionalism makes extended cognition a conceptual possibility, but it doesn't make ExC true. What is needed, additionally, is an account of which functional contributions count as cognitive contributions and which don't. After all, as the critics of ExC have often observed, there undoubtedly will be functional differences between a distributed system and a purely inner system that allegedly realize the same cognitive process. For example, our brain-body-pen-and-paper mathematical system involves visual access to, and the bodily manipulation of, the material symbols on the page. *At some level of functional description*, there will be aspects of these processes that have no counterparts in the purely inner case (e.g. the functions involved in controlling visual gaze). So we will need to know which, if any, of the functional differences in evidence are relevant to determining the cognitive (or otherwise) status of the external elements. In other words, we need to provide a *mark of the cognitive* (Adams and Aizawa, 2008), a scientifically informed account of what it is to be a proper part of a cognitive system that, so as not to beg any questions, is independent of where any candidate element happens to be spatially located (Wheeler, 2010a, 2010b, 2011a, 2011b). The idea is that once a mark of the cognitive is specified, further philosophical and empirical legwork will be required to find out where cognition (so conceived) falls – in the brain, in the non-neural body, in the environment, or, as ExC predicts will sometimes be the case, in a system that extends across all of these aspects of the world.

To see how this might work, let's consider a candidate for a functionalist-friendly mark of the cognitive (Wheeler, 2011a). Newell and Simon famously claimed that a suitably organized "physical symbol system has the necessary and sufficient means for general intelligent action" (Newell and Simon, 1976, p. 116). A physical symbol system (henceforth PSS) is (roughly) a classical computational system instantiated in the physical world, where a classical computational system is (roughly) a system in which atomic symbols are combined and manipulated by structure-sensitive processes in accordance with a language-like combinatorial syntax and semantics. Although Newell and Simon adopted what we might call an unrestricted form of the claim that cognition (which I am taking to be equivalent to "the necessary and sufficient means for general intelligent action") is a matter of classical symbol manipulation, one might reasonably restrict the claim to a narrower range of phenomena, perhaps most obviously to "high-end" achievements such as linguistic behavior, natural deduction and mathematical reasoning. And although classical cognitive scientists in general thought of the symbol systems in question here as being realized inside the head, there is nothing in the basic concept of a PSS that rules out the possibility of extended material implementations (cf. the traditional move of bolting internalism onto functionalism). What this line of reasoning gives us, then, is the claim that "being a suitably organized PSS" is one mark of the cognitive.

So what? Bechtel (1996) defends the view that cognitive achievements such as mathematical reasoning, natural-language processing and natural deduction are (at least sometimes) the result of sensorimotor-mediated interactions between internal connectionist networks and external symbol systems, where the latter (but not the former) feature various forms of combinatorial syntax and semantics. So in these cases the combinatorial structure that, if our mark of the cognitive is correct, is indicative of cognition, resides not in our internal processing engine, but rather in public systems of external representations (e.g. written or spoken language, mathematical notations). The capacity of connectionist networks to recognize and generalize from patterns in bodies of training data (e.g. large numbers of correct derivations in sentential arguments), plus the temporal constraints that characterize real embodied engagements with stretches of external symbol structures (e.g. different parts of the input will be available to the network at different times, due to the restrictions imposed by temporal processing windows) are then harnessed to allow those networks to be appropriately sensitive to the constraints of an external compositional syntax. One might be tempted to conclude from this that a Bechtel-style network-plus-symbol-system architecture is an extended PSS. Of course, more would need to be said before we should give in wholeheartedly to this temptation (Wheeler, 2011a). However, let's assume that any concerns can be addressed. If the further thought that I have flirted with is correct, and being a suitably organized PSS is a mark of the cognitive, then, by virtue of being an extended PSS, the Bechtel architecture is also an extended cognitive system.

There are, then, reasons to think that extended functionalism may deliver ExC. And if extended functionalism is a form of embodied cognition, then the future prospects for embodied cognition are correspondingly rosy. But however revolutionary a result extended functionalism may mandate in relation to where cognition is to be found, the fact is that that outcome neither signals a fundamental change in our understanding of the relationship between cognition and material embodiment, nor (relatedly, as we shall see in a moment) forces us in the direction of a radically new theoretical language in which to carry out cognitive science.

On the first point, consider that, for the functionalist, *and thus for the extended functionalist*, the physical body is relevant "only" as an explanation of how cognitive states and processes are implemented in the material world. Indeed, since multiple realizability, as established by functionalism, requires that a single type of cognitive state or process may enjoy a range of different material instantiations, there is no conceptual room for the specific material embodiment of a particular instantiation to be an essential feature of mind. (Although it may be true that some functions can, as a matter of empirical fact, be implemented only in certain kinds of extant material system, that would be no more than a practical constraint.) So, despite the fact that, along one dimension (see above), extended functionalism is a fully paid-up member of the embodied cognition fraternity, there is another, arguably more fundamental, dimension along which the view emerges as having a peculiarly disembodied character. (Here one is reminded of Searle's observation that functionalism, although standardly depicted as a materialist theory of mind, is fully consistent with substance dualism [Searle, 1980].) This leads us to our second point. In observing that functionalism is the house philosophy of mind in orthodox cognitive science, one thereby highlights the close theoretical connection that obtains between functionalism and representational-computational models of mind. Indeed, the traditional account of the relation between a computational virtual machine and the range of physical machines in which that virtual machine may be instantiated is a particular version of the implementation relation that characterizes functionalist multiple realizability, while the notion of a computational system, via research in AI and computational cognitive psychology, provides a concrete and scientifically productive technological realization of the functionalist schema. It is unsurprising, then, that extended functionalism is routinely pursued in a computational register. Indeed, arguably the most common proposal for an extended cognitive system involves external representational elements that are taken to be constitutive components of the cognitive architecture in question precisely because of the way in which they are poised to contribute to the (distributed) implementation of an information-processing solution to some cognitive problem. Here one might mention the canonical example of the linguistic inscriptions in Otto's notebook that allegedly realize the content of his extended dispositional beliefs (Clark and Chalmers, 1998), but the case above of an extended PSS is an example of the very same signature.

If all this is right, then although extended functionalism, through its rejection of neurocentric internalism, may in some ways advance the embodied cause, it nonetheless leaves the conventional explanatory language of cognitive science fully intact. By analogy with Russian history, this is 1905, not 1917.

The vital body

Alan Turing once remarked that, "[i]n the nervous system chemical phenomena are at least as important as electrical" (Turing, 1950, p. 46). This is a striking comment, especially when one recognizes that, even in the connectionist regions of cognitive science, where the abstract physical organization of the brain, depicted as a distributed network of simple processing units, inspires the abstract structure of the cognitive architecture on offer, neural processes are treated as essentially a matter of electrical signals transmitted along wires. Connectionist networks are, of course, standardly interpreted as computational (and thus as representational) systems, even though they often need to be analyzed in terms of cognitive functions specified at a finer level of grain than those performed by classical computational systems (e.g. using mathematical relations between units that do not respect the boundaries of linguistic or conceptual thought). In the present context, the importance of this interpretation of connectionism is captured by Clark's (1989) philosophical gloss on the field as advocating a "microfunctionalist" account of cognition. As we have seen, functionalism demands an implementational notion of embodiment, a fact that surely remains unaffected by the "micro" nature of the functional grain appropriate to connectionist theorizing. Tying these thoughts together, one might speculate that the conceptualization of the brain as an electrical signaling system – a conceptualization that depends, in part, on abstracting away from the chemical dynamics of the brain – plausibly contributes positively to whatever cogency the computationalist-microfunctionalist picture has. So what happens when the chemical dynamics of brains are brought into view? Does this herald the arrival of a radical notion of embodiment, and thus of embodied cognition, one that brings revolution to the door?

To focus on a concrete example, consider reaction-diffusion (RD) systems. These are distributed chemical mechanisms involving constituents that are (a) transformed into each other by local chemical reactions and (b) spread out in space by diffusion. RD systems plausibly explain the kind of behavior in some unicellular organisms (e.g. slime molds) that researchers in the field of artificial life describe as minimally cognitive, behavior such as distinguishing between different relevant environmental factors, adapting to environmental change, and organizing collective behavior. Many of the molecular pathways present in unicellular organisms have been conserved by evolution to play important roles in animal brains, so an understanding of the ways in which RD systems may generate minimally cognitive behavior will plausibly help us to explain the mechanisms underlying higher-level natural cognition. Against this background, Dale and Husbands (2010) show that a simulated RD system (conceived as a one-dimensional ring of cells within which the concentration of two coupled chemicals changes according to differential equations governing within-cell reactions and between-cell diffusion) is capable of intervening between sensory input (from whiskers) and motor output (wheeled locomotion) to enable a situated robot to achieve the following minimally cognitive behaviors: (i) tracking a falling circle (thus demonstrating orientation); (ii) fixating on a circle as opposed to a diamond (thus demonstrating discrimination); (iii) switching from circle fixation behavior to circle avoidance behavior on the presentation of a particular stimulus (thus demonstrating memory).

To see why this result might be thought to have insurrectionary implications, let's introduce Collins's notion of *embrained knowledge*. According to Collins, knowledge is embrained just when "cognitive abilities have to do with the physical setup of the brain," where the term "physical setup" signals not merely the "way neurons are interconnected," but also factors to do with "the brain as a piece of chemistry or a collection of solid shapes" (Collins, 2000, p. 182). When deployed to generate minimally cognitive behavior, RD systems, characterized as they are by the exploitation of spatio-temporal chemical dynamics, plausibly instantiate such embrained knowledge. But what does this tell us? At first sight it might seem that embrained knowledge must fail to reward any functionalist (or microfunctionalist) gloss. Indeed, given the critical role played by low-level spatio-temporal dynamics in the chemistry of the brain, the notion seems to import a radical understanding of the relation between cognition and physical embodiment, one that Clark (2008a) calls *total implementation sensitivity* and that Wheeler (2010b, 2011a, 2013) calls *vital materiality*. According to this understanding, bodily factors make a special, non-substitutable contribution to cognition, meaning that the multiple realizability of the cognitive, and thus functionalism, must fail.[4]

The preceding line of thought is tempting, but ultimately undermotivated. Indeed, we need to take care not to allow the phrase "have to do with the physical setup of the brain" in Collins's specification of embrained knowledge to run away with us. Multiple realizability does not entail that cognition has *nothing* to do with the physical set-up of the brain. How a function is implemented in a physical system may have all kinds of interesting implications for cognitive science, especially (but not only) in areas such as speed, timing and breakdown profile. For example, let's consider some function that is specified in terms of multiple effects. It may be crucial to understanding the precise temporal structure of those effects that we understand them to be achieved via a form of volume signaling in the brain, in which tiny neuromodulators travel not via neural wiring, but freely diffuse through the brain in clouds, pretty much regardless of the surrounding cellular and membrane structures. (For further information on such chemical signaling, see e.g. Husbands, Smith, Jakobi, and O'Shea, 1998.) All this is consistent with the multiple realizability of the cognitive function in question, even if, in certain circumstances, a different implementation would result in explanatorily significant differences in the temporal structure of thought or behavior. But if "having to do with the physical setup of the brain" does not undermine multiple realizability, then it doesn't necessarily establish the more radical relation between cognition and embodiment that was attracting our revolution-hunting attention.

The sense-making body

Perhaps the seeds of revolution are to be found in a different view of embodiment, one that hails ultimately from contemporary European phenomenology (especially Merleau-Ponty, 1945/1962), but which has had an important influence in some regions of embodied cognition research (e.g. Gallagher, 2005). From this alternative perspective, the body is conceived as the pre-reflective medium by which the world is opened up in lived experience as a domain of value and significance. Crucially, advocates of this style of embodiment (e.g. Dreyfus, 1996) standardly hold that its role in our sense-making practices cannot be understood in representational terms. Given the plausibility of the general thought that representation is necessary for computation, if the sense-making body cannot be understood in representational terms, then neither can it be understood in computational terms. So an embodied cognitive science built on this notion of embodiment would be a revolutionary threat to the received explanatory framework. But how does the challenge to representation get off the ground? Exhibit one in this debate is the so-called *problem of relevance*.

Arguably one of the most remarkable capacities that human beings possess is our fluid and flexible sensitivity to what is, and what is not, contextually relevant in some situation, a capacity which is typically operative, even in the sort of dynamically shifting and open-ended scenarios in which we often find ourselves. Cognitive science ought to explain this capacity, and to do so in a wholly scientific manner (i.e. without appeal to some magical, naturalistically undischarged relevance detector). This is the problem of relevance, also known (in AI) as the frame problem (see e.g. Shanahan, 2009).

If one approaches the problem of relevance from the perspective of orthodox cognitive science, and thus of any view such as extended functionalism that buys into the same fundamental principles, the difficulty manifests itself as the dual problem of how to retrieve just those behavior-guiding internal representations that are contextually appropriate, and how to update those representations in contextually appropriate ways. But how is the computational agent able to locate the relevant, and only the relevant, representations? The natural representationalist thought is that context itself should be modeled. In other words, the agent should deploy context-specifying representations to determine which of her other stored representations are currently relevant. What is wrong with this strategy? According to Dreyfus (1990, with more than a nod to Heidegger), the root problem is this: any attempt to determine the relevance of representational structures using other representational structures is an invitation to an infinite regress, since those latter representational structures will need to have their own contextual relevance specified by further representations. But these new representations will need to have their contextual relevance specified by yet further representations, and so on.

Dreyfus's conclusion is that the problem of relevance is an artifact of representationalism (Dreyfus, 2008, p. 358). So whatever neutralizes that problem must be non-representational in character. One way to unpack this idea is through Merleau-Ponty's (1945/1962) notion of the *intentional arc*, according to which skills are not represented, but are realized as contextually situated solicitations by one's environment (Dreyfus, 2008, p. 340). To borrow an example from Gallagher (2008), when poised to engage in the action of climbing a mountain, the skilled climber does not build a cognitive representation of the mountain and infer from that plus additionally represented knowledge of her own abilities that it is climbable by her. Rather, from a certain distance, in particular visual conditions, the mountain "simply" looks climbable to her. Her climbing know-how is "sedimented" in how the mountain looks to her. The mountain solicits climbing from her.

Rietveld (2012) puts some flesh on this skeleton, by drawing a distinction between different kinds of *affordance* (Gibson's term for the possibilities for action presented by the environment; Gibson, 1979). It is here that the connection with our more radical species of embodiment is finally exposed. Given a specific situation, some affordances are mere possibilities for action, where "mere" signals the fact that although the agent *could* respond to them, such a response would be contextually inappropriate. For example, the table at which I am working affords "dancing on top of," but that possibility is not a feature of my current paper-writing context, so right now I am not primed to respond to it. Some affordances, however, precisely because they are either directly contextually relevant to the task at hand, or have proved to be relevant in similar situations in the past, prime us for action by being what Rietveld calls *bodily potentiating*. It is affordances of the latter sort that are identified by Rietveld as different kinds of Merleau-Pontian solicitation. Figure solicitations are those with which we are actively concerned. For example, in my current paper-writing context, my keyboard summons typing from me. Ground solicitations are those with which we are not currently concerned, but for which we are currently bodily potentiated, and which are thus poised to summon us to act. For example, the teacup on my table that is peripheral with respect to my current focus of activity is nevertheless a feature of my paper-writing context and so is poised to summon me to act in appropriate ways. Human relevance-sensitivity is thus explained by shifting affordance landscapes, by varying patterns of active and body-potentiating solicitations.

Conceived on the Rietveld model, affordances depend constitutively, at least in part, on the kinds of bodies that we have, bodies that can dance, type, grip and lift. So if Dreyfus is right that the skills that explain relevance-sensitivity are realized by our non-representational ability to respond to contextually situated solicitations, and if that non-representational capacity is essentially embodied in the way suggested by Rietveld's affordance-based analysis, then the notion of embodiment that will explain a central feature of cognition is one that will fuel the fires of revolution in cognitive science.

So far so good, but if the present conception of embodiment is to have any traction in cognitive science, its advocates will need to say rather more about the naturalistically acceptable processes that causally explain solicitation and summoning. The most developed story here hails from Dreyfus (2008), who suggests that the solution lies with something like Freeman's neurodynamical framework (Freeman, 2000). According to Freeman, the brain is a non-representational dynamical system primed by past experience to actively pick up and enrich significance, a system whose constantly shifting attractor landscape causally explains how newly encountered significances may interact with existing patterns of inner organization to create new global structures for interpreting and responding to stimuli. Now, it may well be plausible that a Freeman-esque capacity for bottom-up, large-scale, adaptive reconfiguration, avoiding debilitating context-specifying representations, will be part of a naturalistic solution to the problem of relevance. But the fact remains that the all-important holistic reconfigurations of the attractor landscape that are at the heart of things here need to be explained in a way that doesn't smuggle in a magic relevance detector. In relation to this demand, Dreyfus appeals to shifts in attention that are correlated with the pivotal reconfigurations (Dreyfus, 2008, p. 360), but this doesn't seem to do enough work (Wheeler, 2010c). For if the attentional shift is the *cause* of a reconfiguration, then the relevance-sensitivity itself remains unexplained, since shifts in attention are at least sometimes presumably governed by, and so presuppose, sensitivity to what is relevant. But if the attentional shift is an *effect* of the reconfiguration, then we are still owed an explanation of how it is that the relevant attractor is the one that is selected. In sum, the account of relevance-sensitivity on offer from the perspective of Merleau-Pontian, sense-making embodiment may well be revolutionary (non-representational, non-computational) in its implications, but it is dangerously incomplete, because it fails to deliver a compelling causal explanation of the phenomenon at issue. Indeed, the shortfall here is serious enough that one might wonder whether it constitutes a genuine advance over the representationalist alternative.

Breaking the tie

At first sight, the result of our deliberations is an honorable draw, between the reformist embodiment of extended functionalism and the revolutionary embodiment of the sense-making view. After all, neither has a cast-iron solution to the problem of relevance. At this point, however, what ought to kick in is a perfectly healthy methodological principle regarding theory change in science, namely that we should give our support to the competitor theory that requires the less extensive revision of our established explanatory principles, unless we have good reason not to; and that means that, on the strength of the evidence and arguments considered here, the deciding vote goes to extended functionalism, which rejects internalism, while maintaining a conception of the body as an implementational substrate for functionally specified cognitive states and processes that is comfortingly familiar from orthodox representational-computational cognitive science. The right conclusion, then, is that the most plausible of the embodied views that we have canvassed today is, in a theoretically important sense, the least embodied of them all.

Acknowledgements

Some passages in this chapter have been adapted with revision from Wheeler, 2011a, 2011c, 2013; Wheeler and Di Paolo, 2011.

Notes

1 The still-canonical presentation of ExC is by Clark and Chalmers (1998). Clark's own more recent treatment may be found in Clark, 2008b. For a field-defining collection that places the original Clark and Chalmers paper alongside a range of developments, criticisms and defenses of the notion of extended cognition, see Menary, 2010.
2 The case for embedded cognition has been made repeatedly. For just two of the available philosophical treatments, see Clark, 1997, and Wheeler, 2005.
3 In this chapter I assume that the notion of multiple realizability is clear and in good order, but I note that not everyone shares my confidence (see e.g. Shapiro, 2000; Milcowski, 2013).
4 Here I explore only one way in which vital materiality might be motivated. For critical discussion of certain other routes to that position, see Wheeler, 2010b, 2011a, 2013.

References

Adams, F., and Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Blackwell.

Bechtel, W. (1996). What knowledge must be in the head in order to acquire language. In B. Velichkovsky and D. M. Rumbaugh (Eds.), *Communicating meaning: The evolution and development of langauge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge, MA: MIT Press.

——(1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

——(2008a). Pressing the flesh: A tension in the study of the embodied, embedded mind? *Philosophy and Phenomenological Research, 76*(1), 37–59.

——(2008b). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Collins, H. (2000). Four kinds of knowledge, two (or maybe three) kinds of embodiment, and the question of artificial intelligence. In M. Wrathall and J. Malpas (Eds.), *Heidegger, coping and cognitive science: Essays in honor of Hubert L. Dreyfus* (Vol. 2, pp. 179–95). Cambridge, MA: MIT Press.

Dale, K., and Husbands, P. (2010). The evolution of reaction-diffusion controllers for minimally cognitive agents. *Artificial Life, 16*(1), 1–19.

Dreyfus, H. L. (1990). *Being-in-the-world: A commentary on Heidegger's being and time, Division I.* Cambridge, MA: MIT Press.
——(1996). The current relevance of Merleau-Ponty's phenomenology of embodiment. *Electronic Journal of Analytic Philosophy, 4*(3). Retrieved from http://ejap.louisiana.edu/EJAP/1996.spring/dreyfus.1996.spring.html
——, (2008). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian. In P. Husbands, O. Holland, and M. Wheeler (Eds.), *The mechanical mind in history* (pp. 331–71). Cambridge, MA: MIT Press.
Freeman, W. (2000). *How brains make up their minds*. New York: Columbia University Press.
Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.
——(2008). Are minimal representations still representations? *International Journal of Philosophical Studies, 16*(3), 351–69.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Husbands, P., Smith, T., Jakobi, N., and O'Shea, M. (1998). Better living through chemistry: Evolving GasNets for robot control. *Connection Science, 10*, 185–210.
Menary, R. (Ed.) (2010). *The extended mind*. Cambridge, MA: MIT Press.
Merleau-Ponty, M. (1945/1962). *Phenomenology of perception*, London: Routledge.
Milcowski, M. (2013). *Explaining the computational mind*. Cambridge, MA: MIT Press.
Newell, A., and Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery, 19*(3), 113–26.
Rietveld, E. (2012). Context-switching and responsiveness to real relevance. In J. Kiverstein and M. Wheeler (Eds.), *Heidegger and cognitive science* (pp. 105–34). Basingstoke: Palgrave Macmillan.
Rupert, R. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.
Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*(3), 417–57.
Shanahan, M. (2009). The frame problem. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2009 ed.). Retrieved 25 October 2013, from http://plato.stanford.edu/archives/win2009/entries/frame-problem/
Shapiro, L. A. (2000). Multiple realizations. *Journal of Philosophy, 97*(12), 635–54.
Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*, 433–60. Repr. in M. A. Boden (Ed.), *The philosophy of artificial intelligence* (pp. 40–66). Oxford: Oxford University Press, 1990. (Page references to the repr. ed.)
Wheeler, M. (2005). *Reconstructing the cognitive world: The next step*. Cambridge, MA: MIT Press.
——(2010a). In defense of extended functionalism. In R. Menary (Ed.), *The extended mind* (pp. 245–70). Cambridge, MA: MIT Press.
——(2010b). Minds, things, and materiality. In L. Malafouris and C. Renfrew (Eds.), *The cognitive life of things: Recasting the boundaries of the mind* (pp. 29–37). Cambridge: McDonald Archaeological Institute.
——(2010c). Plastic machines: Behavioural diversity and the turing test. *Kybernetes, 39*(3), 466–80.
——(2011a). Embodied cognition and the extended mind. In J. Garvey (Ed.), *The Continuum companion to philosophy of mind* (pp. 220–38). London: Continuum.
——(2011b). In search of clarity about parity. *Philosophical Studies, 152*(3), 417–25.
——(2011c). Thinking beyond the brain: Educating and building from the standpoint of extended cognition. *Computational Culture, 1*. Retrieved from http://computationalculture.net/article/beyond-the-brain
——(2013). What matters: Real bodies and virtual worlds. In I. Harvey, A. Cavoukian, G. Tomko, D. Borrett, H. Kwan, and D. Hatzinakos (Eds.), *SmartData: Privacy meets evolutionary robotics* (pp. 69–80). Berlin: Springer.
Wheeler, M., and Di Paolo, E. (2011). Existentialism and cognitive science. In F. Joseph, J. Reynolds, and A. Woodward (Eds.), *Continuum companion to existentialism* (pp. 241–59). London: Continuum.