

Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind?*

ANDY CLARK
The University of Edinburgh

Mind, it is increasingly fashionable to assert, is an intrinsically embodied and environmentally embedded phenomenon. But there is a potential tension between two strands of thought prominent in this recent literature. One of those strands depicts the body as special, and the fine details of a creature's embodiment as a major constraint on the nature of its mind: a kind of new-wave body-centrism. The other depicts the body as just one element in a kind of equal-partners dance between brain, body and world, with the nature of the mind fixed by the overall balance thus achieved: a kind of extended functionalism (now with an even broader canvas for multiple realizability than ever before). The present paper displays the tension, scouts the space of possible responses, and ends by attempting to specify what the body actually needs to be, given its complex role in these recent debates.

1. Embodiment and Embedding: The Very Ideas

Recent years have seen an explosion of work, both in philosophy and across the many sub-disciplines of Cognitive Science, that is now typically glossed as belonging to the investigation of the mind as 'embodied and environmentally embedded.'¹ The phrase 'mind as embodied and embedded' seems to have been coined by John Haugeland in a similarly

* This project was completed thanks to joint teaching relief provided by the University of Edinburgh and the AHRC Research Leave Scheme (Project: On the Proper Treatment of Embodiment, ID: 112685). Thanks to Michael Wheeler, Jenann Ismael, David Chalmers, Mark Rowlands, and Rob Wilson for useful comments.

¹ A very modest sampling might include Thompson, Varela, and Rosch (1991), Clark (1997), Haugeland (1998), Hurley (1998), Clark and Chalmers (1998), Rowlands (1999), Lakoff and Johnson (1999), O'Regan and Noë (2001), Shapiro (2004), Wilson (2004), Noë (2004), Gallagher (2005). For some critiques, see Adams and Aizawa (2001), Grush (2003). It should go without saying that many if not all of the key ideas have longer and much more varied pedigrees. For a brief survey of these historical roots, see Clark (1997) chapter 8, section 8.

titled paper that was circulating widely in the early 1990's² and that later appeared as Haugeland (1998). There, Haugeland writes that:

If we are to understand mind as the locus of intelligence, we cannot follow Descartes in regarding it as separable in principle from the body and the world...Broader approaches, freed of that prejudicial commitment, can look again at perception and action, at skillful involvement with public equipment and social organization, and see not principled separation but all sorts of close coupling and functional unity...Mind, therefore, is not incidentally but *intimately* embodied and *intimately* embedded in its world. Haugeland (1998) p. 236-237

Intimacy, however, is a famously slippery beast. What claim or claims concerning the role and importance of the body and the environment lie at the centre of recent work on the embodied, embedded mind? In recent years, some of the many projects developed under this broad banner have included: work on externalism and the nature of psychological explanation (Wilson (2004)); work on 'active externalism' and the extended mind (Clark and Chalmers (1998)); work on 'sensorimotor accounts of perception' (O'Regan and Noë (2001)); work on environment-involving accounts of perception, memory, thought and language (Rowlands (1999)); work on the interdependence of conscious perception and action (Hurley (1998)); work on deictic pointers and active vision (Ballard et al. (1997)); and work on the complementarity between biological and technological resources (Clark (2003)). Given this surface diversity, it seems fair to ask what, if anything, forms the deep theoretical core of the embodied, embedded approach?

It helps, at the outset, to put aside some possible readings that are simply trivial or uninteresting. For example, it is obvious enough that the concept of chair in some way reflects facts about the kinds of shape that happen to afford (typical human) sitting. It is also obvious that much daily thought and reason is informed by what we sense and what we do, and that sensings and doings are paradigmatically embodied acts. Creatures with very different bodies to our own won't (*ceteris paribus*- and more on that later) be prone to think quite the same things in the same circumstances as we do. Or take 'knowing how to row'. Knowing how to row, for the average human being, involves knowing how to do certain things with oars and with your arms and legs. What is known is not just how to row, but (if you will) how to row using tools *like this* and a body *like this*. None of this is particularly interesting, either as philosophy or science. The question is whether attention

² I first heard the paper at an Academia Sinica meeting in Taipei in 1993. It appeared in a volume following that meeting (Yu-Houng Houng and Jih-Ching Ho (eds) (1995). The version quoted in the present text is the one found in Haugeland (1998).

to details of embodiment and embedding contributes something important and previously unexpected to an understanding of mind and reason, not whether it contributes at all.

In addition, there is a large complex of traditionally externalist philosophical theorizing that remains orthogonal to the concerns of (most of³) the theorists of embodiment and embedding. Such philosophical externalism concerns the constitutive role that an embedding environment plays in determining the content of some (or perhaps even all) mental states. But that determination is, in a certain sense, merely semantic. In such cases, the external features are thought to play a role in fixing meanings, but do not appear as part of the local computational and bodily mechanisms driving environmentally engaged action. By contrast, most theorists of embodiment and embedding focus on ways in which features of the body, of embodied action, or of the environment may play an active⁴ information-processing role (a clear example is Haugeland's (1998) appeal to the role of the actual road in an agent's 'knowing the way to San Jose').

Having put the trivial and the orthogonal matters aside, what other claim or claims concerning the nature and importance of embodiment and embedding are to be found? Some recent attempts to sort and distinguish various claims and options include M. Wilson's (2002) six-way classification and Shapiro's (2004) three-way division.⁵ I propose, however, to be rather less delicate than either of these and simply to highlight (what seem to me to be) two rather different central claims that organize and inform the bulk of this recent literature.

The first claim is that aspects of body and world can, at times, be proper parts of larger mechanisms whose states and overall operating profile determine (or minimally, help determine) our mental states and properties. Call this the *Larger Mechanism Story (LMS)*.

The second claim is that the specific details of human embodiment make a special and (in a sense to be explored later) ineliminable contribution to our mental states and properties. Call this the *Special Contribution Story (SC)*.

The two stories (I hesitate to call them theses at this early stage of philosophical and scientific investigation) are superficially quite similar. The first, LMS, claims that larger systemic wholes, incorporating

³ An exception to this rule is R.A. Wilson (2004).

⁴ For this reason, Clark and Chalmers (1998) dub their position 'active externalism'.

⁵ Shapiro (2004) distinguishes claims concerning body-specificity, claims concerning the role of language and metaphor, and claims concerning tool-based cognitive extension.

brains, bodies, the motion of sense organs, and (under some conditions) the information-bearing states of non-biological props and aids, may sometimes constitute the *mechanistic supervenience base* for mental states and processes. The second, SC, claims that specific features of the body (and perhaps the world) make a *persistent, non-trivial, and in some sense special* contribution to our mental states and processes. Despite the superficial similarities (each claims that body and world may play a profound role in human cognition) these two stories are both distinct and potentially in some degree of tension, as I'll now try to show.

2. On (Principled) Body-Centrism

Let's take the second story (the one about a special contribution) first. There are various ways to try to make this case, but I'll consider only two. The first appeals to the role of the body in sensing and processing, the second to the role of the body in thought and reason. Shapiro (2004)⁶ lays out both these routes in an especially clear and crisp manner, as part of an argument against what he dubs *the separability thesis* (ST). According to ST a humanlike mind could perfectly well exist in a very nonhumanlike body. Against ST, Shapiro urges us to embrace what he calls *the embodied mind thesis* (EMT) which holds that "minds profoundly reflect the bodies in which they are contained" (op cit p.167).

Why reject ST? One reason, Shapiro tells us, turns on quite basic facts about sensing and processing. Human vision, for example, involves a great deal of sensor movement. We move our heads to gain information about the relative distances of objects, since nearer objects will (courtesy of parallax effects) appear to move the most. Such movements, Shapiro argues, are not simply an aid to vision. They are part and parcel of the visual processing itself. They are "as much a part of vision as the detection of disparity or the calculation of shape from shading" (op cit p.188). Similar points can be made about audition and the placement of the ears on the head. The idea is that:

⁶ Shapiro's larger project is to attempt to undermine the Multiple Realizability Thesis itself, via a two step process. First, he argues that MRT implies that we should not in general be able to predict properties of the brain from properties of the mind. Second, he tries to show that such predictions are in fact possible (for example, there are reasons to think that many details of wiring and placement of neural structures are highly constrained by what those structures are required to do), and thus concludes that MRT is false. The details of these arguments will not concern us today. It is perhaps ironic, though, that at least some of the kinds of story that Shapiro appeals to later in the book, as part of his subsequent argument for the importance of specific bodily form, turn out to exemplify a kind of extended functionalism, and thus threaten to increase, rather than reduce, the opportunities for multiple realizability of mental states.

...psychological processes are incomplete without the body's contributions. Vision for human beings is a process that includes features of the human body...this means that a description of various perceptual capacities cannot maintain body-neutrality and it also means that an organism with a non-human body will have non-human visual and auditory psychologies. Shapiro (2004) p.190

Body-neutrality, for Shapiro, is the idea that “characteristics of bodies make no difference to the kind of mind one possesses” and is associated with the idea that “mind is a program that can be characterized in abstraction from the kind of body/brain that realizes it” (both quotes op cit p.175). Work on the role of bodily movements in visual processing suggests, according to Shapiro, that body-neutrality fails and that human-style vision requires a human-style body.

Another body of research that appears to contest claims of body-neutrality, at least regarding the contents of perceptual awareness, is spearheaded by Alva Noë and Kevin O'Regan, and forms part of the development of an 'enactive' (see Varela et al. (1991)) approach to perception. The central claim is that perception is nothing other than implicit knowledge of so-called 'sensorimotor contingencies'; that is to say, knowledge of the effects of movement on sensory stimulation. The point is not that (trivially) what we see depends on what we do, but that seeing itself is all about our implicit knowledge of the likely effects of our own bodily motions (especially motions of the sensors). Space precludes a full rehearsal of this rich and challenging view, but the conclusions bear repeating:

If perception is in part constituted by our possession and exercise of bodily skills...then it may also depend on our possession of the sorts of bodies that can encompass those skills, for only a creature with such a body could have those skills. To perceive like us, it follows, you must have a body like ours. Noë (2004) p.25

Noë offers an immediate illustration of this claim. As we take off in an airplane, it can look as if the nose lifts up relative to our field of vision. But this is not so, since we are (of course) rising in perfect synchrony with the plane. The nose lifts, but not relative to our visual plane! The illusion occurs because our vestibular system detects the alteration in bodily orientation, and this information impacts our visual experience. According to Noë:

..the example illustrates the way the character of our visual experience depends on our embodiment, that is, on idiosyncratic aspects of our sensory implementation. (op cit p.27)

For another example, this time of a non-illusory nature, my experience of seeing a long straight horizontal line constitutively involves, according to Noë and O'Regan, knowledge of the peculiar sensorimotor contingencies characteristic of my encountering such a line. If I move my gaze smoothly sideways, I will simply encounter more of the line, still in foveal vision. If I move my gaze upwards or downwards however, the line will move out of my foveal region. My experience of seeing the line is, according to Noë and O'Regan, in part determined by my implicit knowledge of these contingencies linking bodily actions to new sensory inputs. Accordingly, a creature with somewhat different bodily and sensory apparatus (a creature with a long thin foveal region punctuated by a central gap, for example) will associate the long straight line with quite different sets of body-related sensorimotor contingencies and hence must, if this theory is correct, enjoy a very different sensory experience (when seeing a long straight horizontal line) to our own.

We shall return to these cases later. For now, I simply note that from the fact that (as seems highly likely) our human experience really does depend in part on many idiosyncratic aspects of our embodiment, it does not follow that *only* a creature thus embodied could have those very experiences. The very most that follows is that, for a creature like us, all other things being equal, *we* would not have that experience were it not for such-and-such an idiosyncratic fact. Other creatures, for all that we have so far had cause to accept, may still have the very same experience courtesy of quite different forms of embodiment and sensory set-up.

Another, very different way of supporting SC (and hence for rejecting ST) appeals to considerations of the role of the body in structuring human concepts. The *locus classicus* here is Lakoff and Johnson's (1980, 1999) work on the role of body-based metaphors in human thought and reason. Many of our basic concepts, they argue, are quite evidently body-based: concepts like front and back, up and down, inside and outside:

If all beings on the planet were uniform stationary spheres floating in some medium and perceiving equally in all directions, they would have no concepts of *front* and *back*. Lakoff and Johnson (1999) p.34

But these basic concepts, they go on to argue, end up structuring our understandings (and our inferences) in more rarefied domains. Happiness and sadness, to take the standard example, are humanly conceived in terms of upness and downness. The specifics of embodiment thus shape the basic concepts that in turn inform (so it is argued) the rest. Summing up the Lakoff and Johnson line, Shapiro (2004) suggests that:

Organisms that didn't have bodies like our own would develop other metaphors to characterize happiness and sadness. *Happy* and *sad* would be structured in other ways and would thus assume different meanings.⁷ Shapiro (2004) p.201

The common upshot of all these arguments, then, is a kind of principled body-centrism, according to which the presence of humanlike minds depends quite directly upon the possession of a humanlike body.⁸

3. Extended Functionalism

It is revealing, I think, that Shapiro's spirited defense of profound bodily⁹ involvement in the mental comes in the larger context of a series of arguments aimed at a different, logically independent but thematically related, target. That target was the thesis of multiple realizability: a staple of non-reductionist Philosophy of Mind ever since the heady days of early Machine Functionalism. At about that time, the notion that minds like ours might be directly identified with their specific *neural* underpinnings was widely cast as a kind of unacceptable meat or species chauvinism, and replaced by the identification of mind as a functional kind, capable in principle of being realized by many different physical substrates (Putnam (1975)—see also Putnam (1960), (Putnam (1967))). In this new regime, mindware stood to neural hardware as software stood to the physical device. Just as the same software could

⁷ That there is something problematic about this argument is evident in the tension between the easy use of a common notion of happiness and sadness in the first quoted sentence and the subsequent assertion that happy and sad would then 'assume different meanings'. But the point, in any case, is simply that arguments stressing the pervasive influence of embodiment on conceptualization look to be arguments for SC, since they assert the ineliminable involvement of bodily details in an account of mental states.

⁸ Recent work on the complex relations between affect and embodiment (Damasio, 1994 & 1999) might seem to establish, rather directly, just such an ineliminable role for the body in a recognizably mental phenomenon. But here too we need to proceed with some caution. For the question, as will become increasingly clear, is not whether or not gross bodily states and processes play a role in the determination of mental states, but whether they play a unique, non-trivial, and ineliminable role. One test for this is to ask whether a creature lacking that kind of body could nonetheless enjoy those very same mental states. Damasio's account is agnostic on this. The body matters, for Damasio, insofar as low-level bodily responses provide a set of 'somatic markers' that afford a kind of compact summary of previous experiences able to impact choice and reasoning in important and often unsuspected ways. Somatic markers thus require there to be some body or other in the loop. But they do not look to require the presence of a body just like ours.

⁹ I use 'bodily' here to refer to the gross physical body, excluding the brain and central nervous system.

turn up on different bedrock machines (as when a PC and a Mac both run Word), so the same kinds of mind might, it was supposed, turn up in various kinds of material form. What mattered was not the bedrock physical forms so much as the abstract patterns of input-to-internal-state-transitions-to-output that the material structures were able to support. Sameness at this rather abstract level was meant to guarantee sameness at the mental level. Or at any rate, any remaining slack was to be taken up by rather arcane details of history and/or distal environmental embedding. As far as the machinery of mind itself was concerned, functional identity fully fixed any contribution to mentality.

Shapiro's appeal to work in embodied, embedded cognitive science depicts it as in spirit rather inimical¹⁰ to the platform-neutral machine functionalist model of mind. But there is a way of understanding the embodied, embedded approach that sees it as extending, rather than undermining, a broadly functionalist story. To bring this possibility into focus, we must next sample the kinds of argument that favour what we earlier dubbed LMS—the *larger mechanism story*.

Arguments in favour of LMS appeal mainly, if not exclusively, to the *computational role* played by certain kinds of non-neural events and processes in online problem-solving. Consider, to pursue an example cited earlier in the text, Ballard et al.'s (1997) account of the use of 'deictic pointers' in a block-copying task. In this task, a subject is given a target pattern of colored blocks that they must copy by moving similar blocks from a reserve area to a new workspace. Using the spare blocks in the reserve area, the subject must recreate the pattern by moving one block at a time from the reserve to the new version they are creating. The task is performed using mouse clicks and a monitor, and the subject's eye motions are constantly tracked. What Ballard et al. found was that repeated rapid saccades to the model were used in the performance of the task: many more than you might expect. For example, the model is consulted both before and after picking up a block, suggesting that when glancing at the model, the subject stores

¹⁰ Inimical to, but not inconsistent with. ST is said to be logically independent of MRT (Multiple Realizability Thesis) since "it is logically possible that a mind could be realized in a number of different kinds of structure, but that all of these structures are contained in similar sorts of bodies (and) it is logically possible that there is only one or a few ways of realizing a humanlike mind but that these few types of realizations can exist in many different sorts of bodies" (Shapiro (2004) p.167). Such concessions make the intended force of the earlier arguments depicting physical structures as proper parts of psychological processes unclear, though Shapiro does add that he is willing to bet that "if there are but a few ways to realize a humanlike mind, probably there are but a few kinds of bodies that could contain such a mind" (op cit p.167).

only one piece of information: either the color or the position of the next block to be copied.

To test this hypothesis, Ballard et al. used a computer program to alter the color of a block while the subject was looking elsewhere. For most of these interventions, subjects did not notice the changes even for blocks and locations that had been visited many times before, or that were the focus of the current action. The explanation was that when glancing at the model, the subject stores only one piece of information: either the color or the position of the next block to be copied (not both). In other words, even when repeated saccades are made to the same site, very minimal information is retained. Instead, repeated fixations provide specific items of information ‘just in time’ for use. The conclusion from this is that:

In the block-copying paradigm... fixation appears to be tightly linked to the underlying processes by marking the location at which information (e.g., color, relative location) is to be acquired, or the location that specifies the target of the hand movement (picking up, putting down). Thus fixation can be seen as binding the value of the variable currently relevant for the task. Ballard et al. (1997) p. 734

Two morals matter for the story at hand. The first is that visual fixation is here playing an identifiable computational role. As the authors (op cit p.725) comment, “changing gaze is analogous to changing the memory reference in a silicon computer.” The second is that repeated saccades to the physical model thus allow the subject to deploy what Ballard et al. dub ‘minimal memory strategies’ to solve the problem. The idea is that the brain creates its programs so as to minimize the amount of working memory that is required, and that eye motions are recruited to place a new piece of information into memory. Indeed, by altering the task demands, Ballard et al. were also able to systematically alter the particular mixes of biological memory and active, embodied retrieval recruited to solve different versions of the problem. They conclude that, in this kind of task at least, “eye movements, head movements, and memory load trade off against each other in a flexible way” (op cit p.732).

The Ballard et al. model is an example of what might be called an extended functionalist approach. It analyses a cognitive task as a sequence of less intelligent sub-tasks, using recognizable computational and information-processing concepts, but applies those concepts within a larger organizational whole. It recognizes the profound contributions that embodiment and environmental embedding make to the solution of the problem, and displays those contributions rather clearly, by identifying the abstract role of specific (both gross-bodily and neural)

operations in real-time performance of the task. The authors are fully aware of this, commenting that their model “strongly suggests a functional view of visual computation where different operations are applied at different stages during a complex task” (op cit p. 735). As a result, a Ballard-style approach is able:

To combine the concept that looking is a form of doing with the claim that vision is computation [integrating the two points by] introducing the idea that eye movements constitute a form of deictic coding...that allow perceivers to exploit the world as a kind of external storage device. Wilson (2004) p.176-177

Bodily actions are thus part of the means by which certain computational and representational operations are implemented. But what makes the cognitive process the one that it is is simply its functional profile (the set of state transitions mediating input and output). The difference is just that this functional profile belongs not to the neural system and its inputs and outputs alone, but to the whole embodied system located in the world.

As a second example of extended functionalism, take Clark and Chalmers (1998) account of the ‘extended mind’. The point of that argument (which I shall not rehearse again today) was to show that external memory traces may sometimes be poised for the control of action in very much the same kind of way as internal memory traces, yielding an extended supervenience base for dispositional beliefs. The claim here was not, implausibly, that an external, passive, encoding might somehow behave exactly like the fluid, automatically responsive resources of internal biological memory. Rather, it was that external encodings were, under certain circumstances, capable of becoming so deeply integrated into online strategies of reasoning and recall as to be only artificially distinguished from proper parts of the cognitive engine itself. The argument thus echoes Ballard et al.’s in depicting a larger integrated system as the extended machinery whose computationally salient states, properties and transformations are supposed to explain specific problem-solving performances.

Or consider an accountant who is extremely good at dealing with long tables of figures. Over the years, she has learnt how to solve specific classes of accounting problem by rapidly scanning the columns, copying some numbers onto a paper scratchpad, then looking to and from those numbers (carefully arrayed on the page) back to the columns of figures. The accountant (let’s call her Ada) does this at lightning speed, and by deploying a variety of minimal memory strategies. Instead of committing complex dependencies to biological memory, Ada follows trails through the numbers, creating external traces every

time an intermediate result is obtained. These traces are in turn visited and re-visited on a 'just in time, need to know' basis, moving specific items of information into and out of short term bio-memory in much the same way as a serial computer shifts information to and from the central registers in the course of carrying out some computation. This process may again be best analysed in extended functional terms, as a set of problem-solving state-transitions whose implementation happens to involve a distributed combination of biological memory, motor actions, external symbolic storage, and just-in-time perceptual access.

R.Wilson's (2004) notions of 'exploitative representation' and 'wide computation' capture the key features nicely. Exploitative representation occurs when a sub-system gets by without encoding some piece of information, in virtue of its ability to track that information in some other way. Wilson gives the example of an odometer that keeps track of how many miles a car has traveled not by first counting wheel rotations then multiplying according to the assumption that each rotation = x meters, but by being built so as to record x meters every time a rotation occurs:

In the first case it encodes a representational assumption and uses this to compute its output. In the second it contains no such encoding but instead uses an existing relationship between its structure and the structure of the world. Wilson (2004) p.163

Wilson's descriptions and central examples can make it seem as if exploitative representation is all about achieving success without representations at all, at least in any robust sense of representation. But this need not be so. Another, very pertinent, range of cases would be those in which a sub-system does not contain within itself a persisting encoding of certain things, but instead leaves those encodings in the world (or in some other sub-system to which it has access). In this kind of case, the larger system may well contain genuine symbolic encodings, consistent with the sub-system merely exploiting (rather than recapitulating) these during problem-solving activity. The case of Ada, described above, would be a case in point. Ada's biological brain does not create and maintain persistent encodings of every figure she generates and offloads onto the page, though it may very well create and maintain persistent encodings of several other key features (for example, some kind of running approximation that acts to check for gross errors). In much the same way as Ballard's block-puzzlers, Ada's biological brain may thus, via the crucial bridging capacities of available embodied action, key its own internal representational and internal computational strategies to the reliable presence of the external

pen-and-paper buffer. Even robustly representational inner goings-on may thus count as exploitative insofar as they merely form one part of a larger, well-balanced process whose cumulative set of state-transitions solves the problem. In this way:

...explicit symbolic structures in a cognizers environment...together with explicit symbolic structures in its head [may] constitute the cognitive system relevant for performing some given task. Wilson (2004) p.184

Our human facility at creating and exploiting external symbolic encodings thus affords a standing invitation to plastic neural wetware to learn and deploy exploitative forms of internal representation and computation in an open-ended range of problem-solving scenarios.

The use of various forms of exploitative representation immediately yields a vision of what Wilson (1994 & 2004) Wilson (2004) dubs ‘wide computationalism’, according to which “at least some of the computational systems that drive cognition reach beyond the limits of the organismic boundary” (op cit p.165). Wide computationalism, stressing as it does the many interactive processes that span brain, body, and world, is also intrinsically dynamics-friendly. Many of the internal representational states invoked will be fleeting, generated on-the-spot, delicately keyed to making the most of other closely coupled internal and external resources. Extended functionalism, of the kind I am describing, is not in any way committed to there being static symbols in the head. On the contrary, it invites us to locate static, classical-looking symbol structures where they belong: out in the world, but making a deep and abiding contribution to online thought and reason nonetheless.

Extended functional systems may thus include motor behaviours as processing devices, and environmental structures as storage and encoding devices.¹¹ Such bodily and worldly elements emerge as genuine parts of extended computational regimes, and are apt for formal description in both dynamical and information-processing terms.¹² The larger

¹¹ Concerning memory, for example, Wilson writes: memory...does not simply stop at the skin but involves engaging with the world through cognitively significant, embodied action...Remembering, on this view, involves exploiting internal, bodily and environmental resources in order to produce some sort of action... Wilson (2004) p.191

¹² The idea that dynamical approaches are incompatible with computational and representational ones is increasingly recognized as a mistake, even by those working at the very heart of the dynamical systems movement: see e.g., Spencer and Schöner (2003). Extended functionalism, of the kind I am describing, is not at all committed to there being static symbols in the head. On the contrary, it invites us to locate static, classical-looking symbol structures where they belong: out in the world, but making a deep and abiding contribution to online thought and reason.

systems thus constituted are, as Wilson insists, unified computational wholes such that “the resulting mind-world computational system itself, and not just the part of it inside the head, is genuinely cognitive” (op cit p.167). Insofar as such a thesis is correct, the cognitive scientist or philosopher of mind who chooses to treat the brain and central nervous system alone as the mechanistic supervenience base for mental states is rather like a neuroscientist who insists that neuroscience proper should not be concerned with the hippocampus or the cerebellum, because (they think) *all the real cognizing goes on in the cortex*, even if (they concede) those other structures sometimes play a role in the transmission and routing of information and control. Extended functionalists¹³ thus reject the image of mind as a kind of input-output sandwich with cognition as the filling (for this picture, and more arguments for its rejection, see Hurley (1998)). Instead, we confront an image of cognition quite literally bleeding into the world.¹⁴

4. Embodiment and Embedding: The Tension Revealed

There is a potential tension, it seems to me, between the kinds of account that typically stress features of human embodiment and the kinds of account that typically stress environmental embedding and intervention. This tension is obscured by the common use of the notion of embodiment, for while embodiment plays a crucial role in each kind of story, the nature of that role is often rather different. Thus, the accounts mentioned in section 2 seem to depict bodily form and sensorimotor patterning as elements that might make a special contribution to human thought and reason. But the accounts mentioned in section 3 seem to depict bodily action and environmental structuring as merely additional elements in a wider computational, dynamical, and representational nexus. What counts for mind, cognition, and mental states, according to these models, is the overall processing economy itself.

To focus the contrast consider Shapiro’s opposition (mounted in the name of Embodied Cognition) to the idea that “the same kind of mind can exist in bodies with very distinct properties” (Shapiro (2004) p.175). On the basis of the kinds of evidence described in section 2 above, Shapiro rejects the idea that “snakelike organisms and creatures of science fiction” (op cit 174) might share our kind of mind. If the

¹³ For example Wilson (2004), Clark (1997), Clark and Chalmers (1998), Dennett (1996), Donald (1991), and Hutchins (1995) all count as extended functionalist treatments according to the present usage.

¹⁴ Despite such claims, Wilson’s account differs from that of Clark and Chalmers in its subsequent treatment of individuals. Minds, Wilson seems to suggest may be wide and extended, while the individual whose mind it is, remains narrow and organismically bound (op cit p. 198).

theorists of embodied cognition are correct, Shapiro suggests, Body Neutrality (the idea that “characteristics of bodies make no difference to the kind of mind one possesses” (op cit p.175)) is false.

It should be clear, however, that something has here gone by a little too swiftly. For imagine now a case in which we have two intelligent beings. One of them is a snake-like creature lying on top of an advanced touch-screen like environment. In this flat-screen setting every little wriggle of the snake can cause specific external symbolic tokens to appear elsewhere on the screen: tokens that are themselves apt for perceptual uptake (perhaps via a kind of Braille). The snake-being (call it Adder) uses this set-up, let us suppose, to carry out the same complex accounting as Ada in our earlier example. According to LMS (the Larger Mechanism Story) there is no reason to suppose (from anything we have said so far) that the accounting-relevant cognitive states of Ada and Adder need differ in any respect. Each implements the same extended computational process, and even (we may suppose) divides the biological and non-biological contributions in the same way.

More radically, but still consistent with LMS, we may even imagine that there are differences at the level of what gets done where. Enter Odder. Odder performs certain computations internally that Ada and Adder both perform using action and perception routines in the non-biological arena. Here too, the LMS theorist, or extended functionalist, is at liberty to believe that the very same cognitive and mental states might be being implemented, with nothing distinguishing the cases apart from some non-essential matters of location. Just as, for the standard approach, we need not care (within sensible limits) exactly where *within* the brain a given operation is performed, so too (it might be urged) we should not care whether, in some extended computational process, a certain operation occurs inside or outside some particular membrane or metabolic boundary.

The deep issue here becomes clearer if we now recall the main way of defending total implementation sensitivity mentioned in the discussion of Noë and O'Regan. The key move there, recall, was to treat the experiential dimensions of perception as a special case. Here, we speculated, it makes sense to think that the full details of sensory implementation play a unique functional role: a role so unique that only a total copy can actually fit the fine-grained functional profile. (This is rather like those job ads that are so well tailored to a favored candidate that no other human being can ever hope to be as good a fit!) What the example of Odder highlights is the difficulty of insulating the perceptual dimensions from all the rest. Odder is by no means a duplicate in the sensory arena, and the distribution of operations varies, between Odder and the others, in ways that criss-cross the perceptual/non-perceptual divide.

Yet there is a strong (full-bloodedly) functionalist intuition that *considered as a form of 'accounting intelligence', Odder, Adder, and Ada form a computational equivalence class.*

The mistake to avoid, it now seems clear, is to move from:

- (1). Bodily structures and worldly interventions can be active and crucial participants in extended information-processing routines

to

- (2). Bodily structures and worldly interventions must in all cases play a special role such that *sameness of mental state requires sameness of bodily structure.*

The spirit of LMS, it seems to me, is precisely to reject (2) by stressing overall systemic integrity at the expense of any special role for the contributions of body, brain, or world. Fans of SC, by contrast, will think that LMS fails to do justice to what matters most about the body, and that it is simply repeating an old functionalist mistake in a new (extended, distributed) setting.

5. Remedies for the Flesh?

There is, it seems, a *prima facie* tension at the heart of the program so easily (so unitarily) glossed as the study of “embodied, embedded cognition.” It is the tension between an extended, situated-reason friendly version of good old fashioned functionalism, and something more fundamentally fleshy: the idea that features of the body make a special and in some sense non-negotiable contribution to mind and mentality.

One quick way to relieve the tension would be to argue that the body makes a special contribution to conscious experience (and perhaps also thus to the bedrock for metaphorical thought) but not to cognition considered more generally. Thus Ada, Adder, and Odder may indeed all implement the very same accounting algorithm, but the differences in embodiment will (on this account) necessarily make a difference at the level of conscious experience. The wild card in this whole debate is thus our old friend phenomenal experience itself. Thus consider Shapiro’s observation that:

The instructions by which the human brain computes relative depth do not work in creatures with eye configurations other than those in a human being. This is the sense in which depth perception is embodied. The procedures by which human beings perceive depth—a fact about

human psychology—are contingent on a fact about human bodies. Shapiro (2004) p.188

Recall that from facts such as these, Shapiro concludes that “human vision needs a human body” (op cit p.189). Such a claim is, however, importantly ambiguous. It might mean only that the brain’s algorithms factor in the bodily structures and opportunities. This is surely correct, and we shall return to it later. It might mean that being able to make the kinds of gross visual discrimination that we can make *requires* having exactly the same kind of body (in respect of eye configuration at least) as we do. But this claim is surely false, since an alternative distribution of the very same information processing steps, in some differently brained and differently bodied being, must be capable of implementing that same algorithm (as LMS is frequently at pains to point out).¹⁵ Or it might (finally) mean that any such alternative implementation need not preserve the qualitative feel of human depth perception: a qualitative feel that is somehow tied not to the abstract algorithm but to the use of two eyes located a certain distance apart. Perhaps, then, the fans of SC are motivated, at least in part, by the intuition that the body plays a special, non-trivial role in the determination of *qualitative experience*.

The quickest reconciliation, then, is to treat the body or sensory apparatus as itself making some special kind of contribution, one that cannot help but impact (in non-trivial ways) certain qualitative aspects of our mental life. This is probably the best way to understand Noë’s assertion that “the character of our experience depends on...idiosyncratic aspects of our sensory implementation” (op cit p.26). If you think that the sensory implementation plays a unique role that contributes directly to experiential content, you may very well think that every difference in implementation makes a real (though perhaps vanishingly small) difference to the felt nature of the experience itself. In this way, you get both to be a kind of functionalist (arguing that it is the role of the physical structures in mediating patterns of sensorimotor contingency that matters) *and* to assert that (for perceptual experience at least) every difference in certain aspects of physical implementation makes a difference.

¹⁵ Thus consider FLICKER. Flicker is a creature with just one eye that moves very rapidly from side to side of its face, sending signals only while at the two locations that happen to match those of the human eyes. With some canny tweaks of the neural control and downstream sensory post-processing circuitry, such a being could implement precisely the same basic depth perception algorithm as ourselves. The situation would be not unlike the use of a fast serial computer to simulate a parallel processing device.

It is by no means obvious, however, that we can stably reconcile any recognizable form of functionalism with such *full-and-principled* sensitivity to all the details of a being's embodiment and/or sensory apparatus. For a broadly functional (or even just a broadly computational and representational) view of the underpinnings of perceptual experience demands, it seems to me, that it be in principle possible that two beings could be different in respect of gross sensory apparatus and embodiment and yet, courtesy perhaps of compensatory differences in key aspects of downstream processing, end up realizing the same set of experience-determining functionally specified state-transitions. Noë (2004) (and also O'Regan and Noë (2001)) seem to leave no room for this even as a bare possibility. Noë is explicit that "to see *as we do*, you must...have a sensory organ and a body like ours" (Noë (2004) p. 112, italics in original).¹⁶

Perhaps this is right, and experience is non-trivially permeated by the full details of biological embodiment. My own view (see e.g. Clark (1999) Clark (2001): see also Jacob and Jeannerod (2003)) is that this is unlikely to be true. Experience, in presenting us with a world fit to engage by action and by reason, need be sensitive only to certain aspects or features of the sensorimotor contingencies our embodiment dictates. By simply identifying experiences with implicit knowledge of the full suite of contingencies defined at the sensorimotor surfaces, the strong sensorimotor account leaves no room for compensatory downstream adjustments to yield identical experiences despite surface dissimilarities.¹⁷ Nor does it leave room for small differences at the sensorimotor surfaces to be such as to make no experiential difference, courtesy of failing to deliver any *salient* differences in signals to downstream processors. Perhaps, that is to say, downstream processing provides a kind of grid relative to which certain differences at the level of the sensory inputs (and associated contingencies) simply fail to *make* a difference. But whatever the cognitive scientific niceties, the point for present purposes is just that no global reconciliation between LMS and

¹⁶ Such an account makes it in principle impossible for a differently embodied being to fully share human perceptual experiences. In this vein, the account has been accused (Clark and Toribio (2001)) of a kind of 'sensorimotor chauvinism'. Noë (2004, p.26-28, and p. 113) rejects the charge of unacceptable chauvinism arguing that while every difference in sensorimotor contingencies must indeed make some slight difference to perceptual experience, there is plenty of room to allow that differently embodied beings may enjoy sufficiently similar experiences for each to count as, say, visual.

¹⁷ Thus Noë (personal communication) does indeed assert that "you couldn't have the very same experience unless you have the same underlying sensorimotor exercise." This may turn out to be true, but it is not yet obvious to me why it must be true, or how we can at this time know it to be true.

the strong sensorimotor model looks likely, since the latter depicts the sensorimotor dynamics alone as fixing, with extreme sensitivity, the nature of our perceptual experience. This forecloses the possibility of the same experience being brought about in some other sensorimotor context in virtue of the details of a larger functional organization (or, contrariwise, of a different perceptual experience being brought about, despite sameness of sensorimotor contingencies, in virtue of some larger functional organization).

A related worry threatens at least the strongest versions of Lakoff and Johnson's claims concerning the tight links between forms of embodiment and basic conceptual repertoires. For what embodied experience actually delivers as the baseline for learning and metaphorical thought surely depends on some complex mixture of bodily form, environmental structure and (possibly innate) downstream internal processing. Here too, compensatory adjustments in either of the two non-bodily arenas look likely to make available forms of thought and reason that are not tethered in any simple way to the gross bodily bedrock.

For these reasons, I think it would be unwise to rest content with the bald assertion that bodily form and structure plays some special role in the determination of conscious experience. Instead, I think we should try to understand the body as playing a distinctive computational role: one that impacts both conscious and non-conscious cognitive strategies, and that explains why the body *matters* without making the body matter mysteriously.

One obvious place to start is by observing that for many (perhaps all) information-processing problems, there will be an elegant, representationally low-cost solution that is in several important senses relative to the gross physical properties of the implementing apparatus. For example, an industrial solution to the problem of fitting together small highly engineered parts may be built around the provision of flexible rubber mountings for the assembler arms. The physical structure here reshapes the information processing required. Subtract the rubber mountings and the best solution involves the repeated computation of multiple visual feedback loops to guide repeated attempts. But with the rubber mountings in place, the system gives along two spatial axes, and even roughly visually guided parts fall into place "just as if millions of tiny feedback adjustments to a rigid system were being continuously computed" (Michie and Johnson, 1984 p. 95).

Similar effects flow from the concrete details of sensor placement. A system with a certain spatial distribution of sensors for heat or light will not need to deploy multiple steps of inference to determine whether certain signature patterns are present or absent. Moreover, the fixed

relations between bodily-mounted sensors obviate the need to constantly determine how input at point X relates to input at point Y. Such relations may be either constant (as between two fixed eyes) or else vary systematically (where X and Y are independently controllable or moveable, as in the case of the left and right index fingers). In either case, the properties of the body keep the sensory inputs in a certain kind of alignment, and this can be simply assumed (rather than explicitly represented) by the algorithms that use the sensory inputs as sources of problem-solving information.

The body is also the point at which willed action, if successful, first impacts the wider world. This sounds trivial, but is actually profoundly important. When conjoined to the observation that, in the typical human case, these points of willed action include all our voluntary sensor movements, it yields the intuitive understanding of the body as the common and persisting locus of sensing and action. Extensive work on the technologies of telepresence¹⁸ suggests that the human sense of presence, of being at a certain place in space, is fully determined by our ability to enter into closed loop interactions, in which willed sensor motions yield new sensory inputs, and by our ability to act upon at least some of the items thus falling within sensory range. The phenomenological sense in which we typically ‘inhabit’ (rather than painstakingly control) our own bodies is determined by the fluency with which we control the sensors and actuators in ways that enable successful interaction to occur.

Finally, the body, by being the immediate locus of willed action, is also the gateway to intelligent offloading. The body is the primary tool for the intelligent use of environmental structure (see Kirsh (1995)) and acts as the mobile bridge that allows us to use the external world in ways that simplify and transform internal problem-solving. The body is thus the go-between linking these two different (internal and external) sets of key information-processing resources. The body’s role in such cases is that of an instrument enabling the emergence of a new kind of information-processing organization. This role may, without too much exaggeration, be likened to that of the corpus callosum. Both are key physical structures whose cognitive role is in part to allow distinct sets of resources to engage in highly integrated forms of problem-solving activity.

In these closing comments I have been speaking as if the body is, just as it happens, the locus of willed action, the point of sensorimotor confluence, the gateway to intelligent offloading, and the stable (though not permanently fixed) platform whose features and relations can be

¹⁸ For a review, see Clark (2003) chapter 4.

relied upon in the computation of certain information-processing solutions. But in fact, I am inclined to go further and to assert not just that this is what the body does, but that this (or some list quite like it) is what, at least for all cognitive scientific purposes, the body *is*. I am inclined, that is, to identify the body with whatever plays these roles in a unified information-processing economy.

The immediate upshot of this is that the body, insofar as it is cognitively significant, turns out to be itself defined by a certain complex functional role. Notice also that nothing here requires a single persisting body in ordinary 3-dimensional space. Instead, there could be genuine but scattered forms of embodiment, embodiment in virtual or mixed realities, and multiple embodiments for a single intelligence (for more on these topics, see Clark (2003), Clark (2007), Ismael (2007)). It is merely a contingent (and increasingly negotiable) fact about human embodiment that the body is both the metabolic centre *and* the bridge to sensory presence and intelligent action. In exotic cases,¹⁹ the metabolic center is detached from the more cognitively important loci for sensing, acting and intelligent offloading: loci that collectively determine our sense of presence in the world.

A deeper resolution of the apparent tension between SC and LMS is now in sight. For the cognitive importance of the body, if this is on the right track, is fully exhausted by its ability to play a certain functional role in an intelligent organization. The distinctiveness and importance of this role is what explains the intuition that the body makes a special and quite pervasive cognitive contribution. But because it *is* nothing but a complex functional role, there is nothing cognitively significant about the bodily contribution that is not fully captured by reflection upon its several (and potentially separable) computational and information-processing contributions. The body *is* special. But we should understand its specialness through the familiar lens of our best information-processing models of mind and cognition.

6. Conclusions: Embodiment, Functionalism, and the Search for a Balanced Mind

There is a certain tension, or so I have argued, between two strands of thought prominent in the recent literature on the embodied, embedded mind. One of those strands depicts the body as intrinsically special, and the details of a creature's embodiment as a major and abiding constraint on the nature of its mind: a kind of new-wave body-centrism. The other depicts the body as just one element in a kind of equal-partners dance between brain, body and world, with the nature of the

¹⁹ For a famous meditation on this theme, see Dennett (1981).

mind fixed by the overall balance thus achieved: a kind of extended functionalism (now with an even broader canvas for multiple realizability than ever before). Where some theorists of embodiment see a unique, non-trivial, and ineliminable role for the body in the determination of mental states, some theorists of embedding thus see only a larger functional whole with considerable slack concerning which operations are performed using which kind of resource.

The proper resolution of this tension, I have argued, is to display the body as (for all cognitive purposes) nothing but the item, or items, that play a certain complex functional role in an information-processing economy. Within such an economy, mental sameness is determined by the overall balance achieved using neural, bodily and environmental resources. The body plays a special role in determining and stabilizing this balance and as such it is a key player on the cognitive stage. But this gives us no reason at all to insist that mental sameness requires gross bodily sameness. To do so would be to insist on a kind of compartmentalization that is inimical to the true cognitive role of the body, which is to act as a bridge enabling biological intelligence and the wider world to intermingle in the service of adaptive success.

References

- Adams, F and Aizawa, K (2001) "The Bounds of Cognition," *Philosophical Psychology* 14:1: 43–64.
- Ballard, D.H., M.M., Hayhoe, P.K., Pook, and R.P.N., Rao (1997) "Deictic codes for the embodiment of cognition," *Behavioral and Brain Sciences* 20 (4): 723–767.
- Clark, A (1997) *Being There: Putting Brain, Body and World Together Again* (MIT Press, Camb. MA).
- (1999) "Visual Awareness and Visuomotor Action" *Journal Of Consciousness Studies* 6:1–18
- (2001). "Visual Experience and Motor Action: Are The Bonds Too Tight?" *Philosophical Review* 110 (4): 495–519.
- (2003) *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence* (Oxford University Press, NY).
- Clark, A (2007) "Reinventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind." *Journal of Philosophy and Medicine* 32:3: 263–282.
- Clark, A. and Chalmers, D. (1998). "The Extended Mind." *Analysis* 58: 10–23 (Reprinted in (P. Grim, ed) *The Philosopher's Annual*, vol XXI, 1998).
- Clark, A and Toribio, J (2001). "Sensorimotor Chauvinism?: Commentary on O'Regan, J.K., and Noë, A. 'A sensorimotor approach to

- vision and visual consciousness’.” *Behavioral and Brain Sciences* 24 (5): 979–980.
- Damasio, A (1994) *Descartes’ Error* (Grosset/Putnam, NY).
- (1999) *The Feeling of What Happens* (Harcourt, Brace and Co, NY).
- Dennett, D (1981) “Where am I?” in D. Dennett, *Brainstorms* (Harvester Press, Sussex).
- (1996) *Kinds of Minds* (Basic Books, NY).
- Donald, M (1991) *Origins of the Modern Mind* (Harvard University Press, Cambridge, MA).
- Gallagher, S (2005) *How the Body Shapes the Mind* (Oxford University Press, Oxford).
- Grush, R (2003) “In Defense of some ‘Cartesian’ Assumptions about the Brain and Its Operation.” *Biology and Philosophy* 18: 53–93.
- Haugeland, J (1998) “Mind Embodied and Embedded” in J. Haugeland *Having Thought: Essays in the Metaphysics of Mind* (Harvard University Press, Cambridge, MA) 207–240. (Originally appeared in *Acta Philosophica Fennica* 58 (1995) p.233-267 (a special issue on Mind and Cognition edited by L. Haaparanta and and S. Heinamaa)).
- Houng, Y and Ho, J (eds) (1995) *Mind and Cognition* (Academia Sinica, Taipei).
- Hurley, S (1998) *Consciousness in Action* (Harvard University Press, Cambridge, MA).
- Hutchins, E (1995) *Cognition in the Wild* (MIT Press, Camb. MA).
- Ismael, J (2007) *The Situated Self* (Oxford University Press, Oxford).
- Jacob, P and Jeannerod, M (2003) *Ways of Seeing: The Scope and Limits of Visual Cognition* (Oxford University Press, Oxford).
- Kirsh, D (1995) “The Intelligent Use of Space.” *Artificial Intelligence* 73(1-2): 31–68.
- Lakoff, G and Johnson, M (1980) *Metaphors We Live By* (University of Chicago Press, Chicago).
- (1999) *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought* (Basic Books, NY).
- Michie, D and Johnson, R (1984) *The Creative Computer* (Penguin, NY).
- Milner, A and Goodale, M (1995) *The Visual Brain in Action* (Oxford University Press, Oxford, UK).
- Noë, A (2004) *Action in Perception* (MIT Press, Camb. MA).
- O’Regan, J.K. and Noë, A (2001) “A sensorimotor approach to vision and visual consciousness.” *Behavioral and Brain Sciences* 24(5): 939–973.
- Parfit, D (1984) *Reasons and Persons* (Oxford University Press, Oxford).

- Pettit, P (2003). “Looks as Powers.” *Philosophical Issues* 13: 221–252.
- Pitcher, G (1970) *A Theory of Perception* (Princeton Univ Press, Princeton, N.J.).
- Putnam, H (1960) “Minds and Machines” in S. Hook (ed) *Dimensions of Mind* (New York University Press, NY).
- Putnam, H (1967) “Psychological Predicates” in W. Capitan and D. Merrill (eds) *Art, Mind and Religion* (Pittsburgh University Press, Pittsburgh, PA) 37–48.
- (1975) “Philosophy and our Mental Life” in H. Putnam (ed) *Mind, Language, and Reality* (Cambridge University Press, Cambridge, UK) 291–303.
- Rowlands, M (1999) *The Body in Mind: Understanding Cognitive Processes* (Cambridge University Press, Cambridge, UK).
- Shapiro, L (2004) *The Mind Incarnate* (MIT Press, Camb. MA).
- Spencer, J and Schöner, G (2003) “Bridging the representational gap in the dynamic systems approach to development” *Developmental Science* 6: 392–412.
- Taylor, G (1979). “The Validity of Transcendental Arguments.” *Proceedings of the Aristotelian Society* 79: 151–165.
- Varela, F., Thompson, E. and Rosch., E (1991) *The Embodied Mind* (MIT Press, Camb. MA).
- Wilson, M (2002) “Six views of embodied cognition” *Psychonomic Bulletin and Review* 9(4):625–636.
- Wilson, R. A. 1994. “Wide Computationalism”. *Mind* 103: 351–372.
- (2004) *Boundaries of the Mind: The Individual in the Fragile Sciences—Cognition* (Cambridge University Press, Cambridge, UK).