

# The Bodily Formats Approach to Embodied Cognition

ALVIN I. GOLDMAN

## Chapter Overview

In the past few decades, many practitioners of cognitive science and philosophy of mind have staked out programs and positions under the label of “embodied cognition” (EC). They have widely differing views, however, of what embodiment consists in and why a program of embodied cognition might be an improvement over classical cognitivism. Frederique de Vignemont and I have recently proposed a general characterization of embodied cognition (Goldman & Vignemont, 2009), and I have expanded on the ramifications of this conception when one adds to it evolutionary considerations and a certain attractive conception of neuroarchitecture (Goldman, 2012). On this occasion, I review the definitional proposal in the context of other conceptions of embodiment and show how an assortment of empirical evidence lends strength to our proposal. Finally, I reply to Lawrence Shapiro (this volume, Chapter 3), who raises a series of challenges for our conception.

Existing proposals for EC can be divided into two general categories: proposals predominantly derived from computer science, artificial intelligence (AI), and robotics; and proposals predominantly derived from cognitive psychology and cognitive neuroscience. In the first section, I look at samples of both kinds of proposals. The second section then reviews the definition of embodiment we have proposed and how extensive the range of embodiment is likely to be given the empirical findings that have already emerged. Finally, in the third section, I reply to Shapiro’s comments.

## Embodied Cognition: Highlights of Alternative Conceptions

A chief impetus for the embodied cognition approach in AI—as Michael Anderson (2003) tells the story—is in terms of a reaction by selected AI practitioners to good old-fashioned artificial intelligence (GOFAI). As Anderson explains, the story is usefully begun with Descartes’s philosophy. Descartes drew a sharp distinction between animals and humans. He regarded animals as mere automata, having sensations but no thought or language. True intelligence is to be identified with higher-order reason and language. These ideas were revived in the twentieth century by GOFAI and the classical cognitivist movement in general. EC in computer science and AI is substantially a reaction to these ideas.

A fundamental feature of classical cognitivism is the thesis that thought is fundamentally the manipulation of abstract symbols in accord with explicit rules. In contrast to this high-level or top-down approach to intelligence, Rodney Brooks (1999) has advocated an approach to intelligence that proceeds from the bottom up, and specifically urges us to recall our evolutionary heritage. Human beings are largely continuous with our forebears, from whom we inherited a substrate of systems for coping with the environment. Brooks therefore advocates the following, decidedly un-Cartesian idea: “The study of that [inherited] substrate may well provide constraints on how higher-level thought in humans could be organized” (1999: 135). Brooks argues that representation is the wrong unit of abstraction for building the bulkiest parts of intelligent systems. A substrate of perceptual and behavioral capacities must be established in order to *ground*, or give meaning to, any and all mental symbols. Ungrounded abstract symbols cannot by themselves constitute intelligence.

These ideas echo themes from several philosophers, such as Heidegger (1962) and Merleau-Ponty (2002), whose work is widely cited in the EC literature. Heidegger contended that “being in the world”—involving practical agency and interactive coping—is essential to intelligence and mindedness. Merleau-Ponty argued that representations are “sublimations” of bodily experience, and the employment of such representations “is controlled by the acting body itself, by an ‘I can’, not an ‘I think that’ ” (as explained by Hilditch, 1995: 108–109).

Unfortunately, the precise role of the body in this literature is mainly gestured at rather than clearly delineated. Moreover, it is difficult to pinpoint the empirical support for the theses that one can sink one’s teeth into. These are among the principal reasons why I find them less satisfactory or persuasive than the styles of support for EC to be found in other branches of cognitive science, specifically cognitive psychology and neuroscience. Particular strands of research in these fields disclose much more specificity vis-à-vis phenomena that are crucial to EC and its empirical viability. However, Brooks’s emphasis on our evolutionary heritage and the low-level substrate it leaves with us is an emphasis prominently retained in the positive web I shall weave.

Another important theme that carries over from the AI conception of EC is the idea that thought and language must be grounded in low-level cognition, specifically, in sensorimotor cognition (which is commonly equated with embodied cognitions). The traditional idea of classical cognitivism is that pure, amodal cognition occupies a level of cognition entirely segregated from perception and motor execution. The latter types of cognition are executed via an assortment of special-purpose modules that are encapsulated from information in higher-level cognition (Fodor, 1983). Such amodal cognition, or symbol systems, does not need—or cannot get—grounding in modal cognition. Embodiment-leaning researchers, by contrast, adduce evidence that purports to show that high-level cognition—even language (generally assumed to be at the apex of cognitive capacities)—is deeply interwoven with sensorimotor cognition. Even the semantic content of verbs referring to bodily actions are said to be understood—at least partly—in terms of modal cognitions. This provides alleged support for an embodied grounding thesis. Whether it is adequate support, and whether semantic grounding should be considered the crucial test of embodiment, remains to be seen.

There are two entirely different ways of formulating and/or interpreting EC theses. On one interpretation, the *body itself* (and its various parts) plays a crucial role in cognition, a much more pervasive role than classical cognitivism recognizes. On the second interpretation, it is *representations* of the body and its parts that are so pervasive and important to cognition. Theorists like Brooks (1999), Thelen and Smith (1994), and others contend that cognition is significantly mediated by the body's interaction with its environment, where this interaction does not take the form of the mind's *representation* of the body.<sup>1</sup> This sort of thesis lies at the heart of the nonrepresentationalist form of EC. On the other side stands a large and growing group of investigators who focus on the prevalence in cognition of representations of the body's condition and activity. This important difference sometimes goes unremarked, whereas I take it to be critical to a clear elucidation of the kind of approach to EC one means to develop.<sup>2</sup>

Consider the work of George Lakoff and Mark Johnson (Lakoff, 1987; Lakoff & Johnson, 1999). Their dominant theme is the pervasive use of body-related metaphor in language and thought. The core idea here is the intensive use of representations of (parts of) the body in metaphor and associated uses of language. In *Philosophy in the Flesh*, for example, they describe how we use bodily “projection”:

Bodily projections are especially clear instances of the way our bodies shape conceptual structure. Consider examples such as in front of and in back of. The most central senses of these terms have to do with the body. We have inherent front and backs. We see from the front, normally

move in the direction the front faces, and interact with objects and other people at our fronts . . .

We project fronts and backs onto objects. What we understand as the front of a stationary artifact, like a TV or a computer or a stove, is the side we normally interact with using our fronts. (Lakoff & Johnson, 1999: 34)

Although the beginning of this passage speaks of *bodies* (per se) as shaping our conceptual structure, the core message of the passage is that our *conceptual structure* is pervaded by representations of our bodies. Our conceptualization of objects is dominated by how we conceptually relate other objects to our own bodies. This is all, fundamentally, about representation. Similarly, Lakoff and Johnson's talk of "projecting" the body onto other objects is really addressed to how we *think* about other objects—namely, by conceptually relating them to representations of our own body.

The centrality of representation is also transparent in many EC arguments that address the grounding problem. Lawrence Barsalou is a leading proponent of this theme (1999; see also Prinz, 2002). He develops the idea that abstract thought—generally referred to as *amodal* symbols—is grounded in the experience of perceptual, motor, and other forms of nonabstract thought—referred to as *modal* symbols. Indeed, amodal thought does not merely originate in modal thought, according to Barsalou, but ultimately reduces to *simulations* of (i.e., the revival and re-cycling of) the same modal experiences that previously occurred during perception. Since all modal experience is subsumed under the heading of "embodied," it emerges that all amodal thought is also embodied. This, of course, contrasts sharply with classical cognitivism. Here is an example of how Barsalou (2008) expounds these ideas. First, the classicist account of higher-order thought in language comprehension is rendered as follows:

During language comprehension, hearing the word for a category (e.g., "dog") activates amodal symbols transduced from modal states on previous occasions. Subsequent cognitive operations on category knowledge, such as inference, are assumed to operate on these symbols. Note that none of the modal states originally active when amodal symbols were transduced . . . are active during knowledge representation. . . . Instead, amodal symbols are assumed to be sufficient, with modal states being irrelevant. (Barsalou, 2008: 12)

His own preferred approach—a grounded cognition approach—is then explained as follows:

On experiencing a member of a category (e.g., "dogs"), modal states are again represented as activations in the visual system, auditory system,

motor system, somatosensory system, and so on. . . . Higher-order cross-modal associations then integrate conjunctive neurons in lower-order association areas to establish a multimodal representation of the experience. . . . [H]earing the word for a category (e.g., “dog”) activates conjunctive neurons in higher-order cross-modal association areas that have previously encoded experiences of the respective category. In turn, these conjunctive neurons activate lower-order conjunctive neurons that partially reactivate modal states experienced previously for the category. These neural reenactments attempt to simulate the modal states likely to occur when actually encountering category members. . . . [T]hese reenactments[are] referred to as *simulations*, given that they result from the brain attempting to simulate previous experience. (2008: 13)

The heavy appeal to simulation supports an embodiment thesis not because simulation per se is associated with embodiment, but because what is simulated is *bodily* experience.

Much of Barsalou’s writing is heavy on the theoretical side. The empirical side is not neglected, but this work is less striking. For more striking experimental findings in support of the theory that higher-order thought is grounded in lower-level cognition, consider the neuroimaging findings of Friedemann Pulvermuller and colleagues (Hauk, Johnsrude, & Pulvermuller, 2004; Pulvermuller, 2005). Pulvermuller and colleagues conducted experiments concerning the link between language comprehension and activation of cortical areas dedicated to action. Traditionally, cortical systems for language and for action control were thought to be paradigms of independent and autonomous functional systems or modules. These systems have different cortical bases in circumscribed areas: motor and premotor cortex in the case of action and left perisylvian areas in the case of language. They are fully dissociable by neurological disease—paralysis and apraxic action deficits versus aphasic language deficits. And they can themselves be subdivided into finer functional subsystems—subsystems for movement of different body parts in the case of action systems and subsystems for speech production versus comprehension, or phonology versus syntax versus semantics, in the case of language (Pulvermuller, 2005: 576). Traditionally, a strict modular organization of language and the action systems, respectively, was supported by the inability of patients who have had a stroke to move one extremity while all other motor and language functions remain relatively intact, or the predominant loss of usage of one category of words.

Modern theoretical perspectives, however, offer a different view, says Pulvermuller. Cortical functions might be served by distributed interactive functional systems rather than by local encapsulated modules. Many links have been demonstrated between the premotor and language areas where they are adjoined, in the inferior frontal cortex, and through long-distance cortico-cortical connections (see Figure 4.1). There are multiple links between the superior temporal

language areas and the motor system, for example, rendering information flow possible between the cortical systems for language and those for action.

Furthermore, Pulvermuller reports, the motor cortex has a somatotopic organization, with the mouth and articulators represented close to the Sylvian fissure, the arms and hands at dorsolateral sites, and the feet and legs projected to the vertex and interhemispheric sulcus (see Figure 4.2). This semantic somatotopy model of action words implies that there are differently distributed networks for the English words *lick*, *pick*, and *kick*. Crucial predictions about the semantic somatotopy model include the following. First, the perception of spoken and written action words should activate cortical areas involved in action and execution in a category-specific somatotopic fashion. Second, the spread of activation is fast, so that specific sensorimotor areas should be activated early in the course of spoken and written word comprehension. Third, activation of the sensorimotor cortex should not require people to attend to language stimuli but should instead be automatic.

Functional imaging experiments in Pulvermuller's laboratory have provided ample support to all of these predictions. Hauk et al. (2004) reported that when participants were instructed to silently read action words that related to the face, arm, and leg, a predicted somatotopic pattern of activation emerged

*Figure 4.1* Connections between the language and action systems. The arrows indicate long-distance cortico-cortical links.

along the motor strip. A similar experiment was carried out with action words embedded in sentences. For example, participants heard action descriptions such as “the boy kicked the ball” or “the man wrote the letter” while their brain was imaged. Specific premotor areas reflecting the different involvement of body part information were found to be active. Hearing different sentences involving *lick*, *pick*, and *kick* activated motor areas that control the tongue, the fingers, and the leg, respectively. These striking findings corroborate earlier findings by Martin, Wiggs, Ungerleider, and Haxby (1996) that the processing of action-related words correlates with the activation of premotor cortex.

Pulvermuller concedes that even if action word processing activates the motor system in a somatotopic fashion, this does not necessarily imply that the motor and premotor cortex influence the processing of action words (2005: 579). But further studies show, he says, that this somatotopy reflects referential word meaning. He concludes as follows:

Action meaning seems to be not only necessary, but also highly relevant for language. Verbs form the grammatical backbone of sentences, and the majority explicitly refer to actions. Tool words, for example, relate to actions for which the tools are made, and words that denote internal states, such as “pain” or “disgust,” can be understood only because both speaker and listener can relate them to similar motor programs that are, by genetic endowment, associated with the expression of pain or disgust. Understanding language means relating language to one’s own actions, possibly because the automatic and extremely rapid linkage of sensory and motor information in our brains benefits comprehension and learning processes. (2005: 661)

All of this lends at least *prima facie* support to the notion that higher-level thought is grounded in low-level representations of motor actions in the

*Figure 4.2* Somatotopy of the motor and premotor cortex: the approximate location of the face/articulators, arm/hand and foot/leg representations.

motor or premotor cortices. However, this support for the language-grounding hypothesis has not escaped criticism by other cognitive neuroscientists. Mahon and Caramazza (2008) take issue with the contention that Pulvermuller's findings establish that "conceptual content is reductively constituted by information that is represented within the sensory and motor systems—the embodied cognition hypothesis" (2008: 59). Mahon and Caramazza go on to equate the grounding theses with the *embodied cognition* hypothesis as a whole. "According to the embodied cognition hypothesis, understanding is sensory and motor simulation" (2008: 59). This hypothesis is contrasted with what they call the "disembodied cognition hypothesis"; namely, "conceptual representations are 'symbolic' and 'abstract', and as such, qualitatively distinct and entirely separated from sensory and motor information" (2008: 59). In effect, then, they see the debate over the truth or falsity of the grounding thesis as equivalent to the debate between the embodied and disembodied cognition approaches, where the latter seems to be equated with classical cognitivism. Mahon and Caramazza side with classical cognitivism:

Concepts of concrete objects (e.g., HAMMER) could plausibly include, in a constitutive way, sensory and motor information. But consider concepts such as JUSTICE, ENTROPY, BEAUTY or PATIENCE. For abstract concepts there is no sensory or motor information that could correspond in any reliable or direct way to their "meaning." The possible scope of the embodied cognition framework is thus sharply limited up front: at best, it is a partial theory of concepts since it would be silent about the great majority of the concepts that we have. (2008: 60)

Mahon and Caramazza also offer additional rebuttals of the embodied cognition (i.e., grounding) thesis as advanced by Pulvermuller and colleagues. Quoting directly from Pulvermuller, they stress the fact that the somatotopic activation of the motor system according to the meaning of action words, while interesting in its own right, does not resolve the issue of whether meaning is embodied. This is because "it is unknown whether the motor system becomes activated *prior to*, or rather only *subsequent to*, access to an 'abstract' conceptual representation" (2008: 62). If the motor system becomes activated only subsequent to an abstract conceptual representation, it may be that the motor system's activation isn't crucially constitutive to the meaning of the concept. That meaning might be supplied by the abstract conceptual representation, not by the motor activations. Addressing the issue in connection with another experimental finding—namely, the action-sentence compatibility effect (Glenberg & Kaschak, 2002)—Mahon and Caramazza make the same sort of point. "According to the embodied cognition hypothesis, the motor system is activated because that activation is causally involved in the semantic analysis of the sentence. According to the disembodied cognition hypothesis, the observed motor activation is due to information

spreading throughout the system” (2008:63). The latter possibility—which has not been excluded—does not imply that the motor system is helping to supply the meaning or content of the concepts in question. So the findings of motor activation do not settle the question in favor of an embodied cognition hypothesis.

I conclude this section with the following summary. First, there are interesting and challenging empirical findings (only a tiny sampling of which have been reported here) that are highly congenial to some forms of embodied cognition. However, depending on exactly how an EC thesis is formulated, this evidence may or may not be close to *establishing* such a thesis. It is extremely important to be quite clear about what, exactly, the EC thesis is supposed to be. Of course, we cannot expect a unique meaning of an EC thesis to be dropped to us from heaven. We must consider definitional options for embodied cognition, and once we choose a particular definition for consideration, we should revisit the question of how strong or weak that evidence is (given the definition). This is exactly what will be undertaken in the next section.

### **A New Approach to Embodied Cognition: The Bodily Formats Approach**

The first task is to specify the kind of definitional question that I think needs to be addressed by EC theorists. How shall we even formulate the problem? We might start with the question, “Under what conditions would it be true to say that cognition is embodied?” This formulation appears to assume, however, that the question is whether cognition *as a whole* is embodied or not. An alternative possibility is to hold that *parts* of cognition are embodied and other parts are not; in other words, that some *tokens* of cognition are embodied and other tokens are not. In that case, we should ask the question, “What property of a token cognition C renders it embodied?” In other words, what is *constitutively* necessary and sufficient for a token cognition to qualify as embodied? This is the “analytical” or “definitional” question I shall initially address.

For reasons already sketched, I prefer a representationalist approach to this question. But a simple-minded representationalist definition is clearly inadequate. Here is what I mean by a “simple-minded” representationalist definition:

(SMR) Cognition (token) C is embodied if and only if C represents a body or part of a body.

This cannot be what EC theorists are after. When a perception or thought is about *another* person’s body, this is not a sufficient reason to view it as embodied. The proposal should at least be amended as follows:

(SMR’) Cognition (token) C is embodied if and only if C represents the cognizer’s *own body* or part of his/her body.

But this too seems on the wrong track. “Outer” senses like vision and hearing can be used to perceive one’s own body; one can see one’s own arm or hear one’s own vocalization. But such perceptions do not smack of embodiment. It is only the use of “inner” senses, or systems of inner bodily monitoring, which introduces the notion of embodiment (at least on the approach we favor). There is ample reason, moreover, to suspect that we share such systems with our animal cousins, as part of our evolutionary heritage. For example, the Parma group that discovered mirror neurons in monkeys began by identifying a “motor vocabulary” in monkey premotor cortex, where individual cells or populations of cells code for particular hand actions such as holding, grasping, breaking, and so forth. (Rizzolatti & Sinigaglia, 2006). In other words, monkeys have a system for internally representing their own movements, just as we do. Systems of bodily representation of these “inner” kinds are what we regard as the key to embodied cognition. Specifically, embodied cognition is the application of *special* systems, systems *dedicated* to inner bodily representation. The study of such systems—not always characterized in these terms—has become a major growth area of cognitive science in the last several decades. These new developments include an impressive array of cases in which systems for bodily representation are also utilized for tasks that go considerably beyond the representation of the body *simpliciter*. Thus, they may provide a *substructure* for higher-level cognitive activities. This is what Pulvermuller’s findings, for example, seem to suggest. Whether these findings support a “grounding” conclusion in the sense of a semantic reduction of linguistic terms to bodily activity terms is a side issue. This question should not usurp attention from other legitimate questions about the organization of cognition.

In our 2009 article, de Vignemont and I drew attention to the (moderate) popularity within cognitive science of the notion of mental *codes* or *formats*. Within a modularist framework (which we did not embrace), there is the idea that each module has its own proprietary format, which may consist in a distinctive vocabulary, syntax, and/or set of computational procedures. Such formats might also have a distinctive array of *contents*, arising from the basic function the code is called on to play. If one disapproves of the “code” or “format” terminology, as being too language-like (which is not intended in the present context), one might speak instead—as I did previously—of representational *systems*. In this chapter, I shall use the terms *format* and (representational) *system* approximately synonymously.

Helping ourselves, then, to the notion of multiple formats or representational systems in the brain, we further postulate that a subset of these systems are (originally) dedicated to representing *bodily* subject matters, in particular, representing bodily states and bodily activity from an *internal* point of view. These are representations of “inner sense” rather than “outer sense.” Now, most bodily representations (and formats of representation) involve descriptive contents. That is, they have contents like “Area A of my body is

currently in state  $\Sigma$ , or is undergoing change  $G$ .” But some classes of representation have imperatival, or “instructional,” contents such as “Effector E: move to the left,” or “Effector F: curl.” Motoric areas, in particular, have representational contents of the imperatival kind. Thus, the premotor and motor areas discussed by Pulvermuller utilize bodily codes the primary function of which is to send messages with imperatival contents. Some of these areas send messages to the mouth or tongue, some to the fingers and arms, and some to the legs.<sup>3</sup>

Many other systems are dedicated to various other body-oriented topics. The primary somatosensory cortex is dedicated to representing the condition of (all parts of) the surface of the body. The so-called pain matrix is a complex system consisting of two functionally specialized networks (see de Vignemont & Jacob, 2012). The sensory-discriminative component represents the intensity of pain and its bodily location. The affective component represents the unpleasantness of a painful experience. It recruits the anterior insula, the anterior cingulate cortex, the thalamus, and the brain stem. In relatively recent work, Craig (2002) explores a system of representation of the entire body that he calls “interoception” (a distinct species of inner sense). This system, the lamina I spinothalamocortical system, conveys signals from small-diameter primary afferents that represent the physiological status of all bodily tissues. Lamina I neurons project to the posterior part of the ventromedial nucleus, or VMpo. Craig calls the VMpo “interoceptive cortex” and argues that it contains representations of distinct, highly resolved sensations, including different types of pain, tickle, temperature, itch, muscular and visceral sensations, and sensual touch.

The most interesting part of the EC story I wish to tell, however, does not reside in the primary functions of these body-representing systems. Rather, it resides in the ways such systems are exploited for *other* cognitive tasks. A theoretical background for understanding these exploitative developments is presented by Anderson (2007, 2008, 2010). Anderson presents a principle of the mind/brain that he called the “massive redeployment hypothesis.” The underlying idea is that, over evolutionary time, or even in ontogeny, neural circuits originally established for one purpose are exapted, recycled, and redeployed for different uses, without necessarily losing their original function. Evidence for this thesis arises from the fact that neural structures we can study now are activated by different tasks across multiple cognitive domains. For example, Broca’s area is not only involved in language processing but also in action-related and imagery-related tasks such as movement preparation, action sequencing, and action recognition. In other words, different cognitive functions are supported by putting many of the same neural circuits together in different arrangements. This is the consequence (largely) of evolution, in which the reuse of existing components for new tasks is favored over the more “expensive” development of additional circuits *de novo*.<sup>4</sup>

Anderson (2010) reviews numerous examples of the massive redeployment hypothesis (some of which I review in Goldman [2012]). Such redeployments can be expected to result in the reuse of bodily formats—originally dedicated to ancient tasks—in the execution of new tasks, for example, the use of motoric representations for language comprehension. Pulvermuller’s identification of a large circuit running from language areas to the motor and premotor cortices is an excellent example of the redeployment of an older (motoric) system, featuring a bodily format, to help execute tasks of language comprehension. Exactly what this “help” consists in remains controversial. But it looks like a lovely example of the reuse of bodily formatted circuits for novel purposes.

Let us now return to the task of providing a definition of an embodied cognition (which was dropped abruptly earlier in the section). I now propose the following. If a cognition C uses an internal bodily format in the process of executing some cognitive task T, then even if task T is in no recognizable sense a bodily task (but rather a higher-level task of some kind), C still qualifies as an embodied cognition. In other words, our proposed definition of an embodied cognition is a B-format linked definition:

(BFC) Cognition (token) C is a specimen of embodied cognition if and only if C uses some (internal) *bodily format* to help execute a cognitive task (whatever the task may be).<sup>5</sup>

This definition alone, of course, does not speak to the question of how extensive a part of human cognition is embodied. However, the experimental literature is chock full of cases of bodily codes being redeployed for non-bodily tasks. In the action-sentence compatibility task, for instance, Glenberg and Kaschak (2002) found that it took longer to respond to a sentence that makes sense when the action described runs counter to the required response motion. So, the simple comprehension of a sentence apparently activated action-related representations. A second category of examples is the reuse of motor control circuits for memory. Casasanto and Dykstra (2010) found bidirectional influence between motor control and autobiographical memory. Participants retrieved more memories and moved marbles more quickly when the direction of movement was congruent with the valence of the memory (upward for positive memories and downward for negative memories.). Many of the much-discussed findings in the mirror-neuron literature are also cases of applying motor codes to nonstandard tasks, for example, not using them to guide one’s own actions but to represent the actions of others. Combining this wide-ranging literature with the definition provided previously makes for a robust case of embodied cognition (even if not a “totalizing” conclusion to the effect that *all* cognition is embodied).

### **Shapiro on the Bodily Formats Approach: Some Replies**

Lawrence Shapiro (this volume, Chapter 3) provides an instructive overview of the state-of-play in research on embodied cognition, which also serves as illuminating background to the considerations I had in mind in advancing a new approach to the subject (Goldman, 2012; see also Goldman & de Vignemont, 2009). I particularly appreciate his framing of my proposal in the context of the history of science. However, although Shapiro has understood and expressed my aims with considerable accuracy, his articulation of them is a bit off track in one important respect. I did not mean to formulate an approach to EC that meshes with the favored approaches of *all* players in the field and could therefore be expected to appeal to them all. My aim was less ecumenical than Shapiro supposes.

Shapiro characterizes my proposal as an “effort to look for a consensus within Embodied Cognition” (this volume, Chapter 3), in order to exemplify the situation in many branches of science where practitioners agree on which are the fundamental problems and assumptions of their field. Shapiro finds such consensus lacking in the current state of play in EC. He sees me as offering a proposal intended to attract the “crowd” of EC enthusiasts and thereby precipitate the desired consensus. He fears, however, that my aim will fall short. But hold on! I never entertained this aim. I am fully aware that my proposal will leave many self-styled EC theorists unmoved and unimpressed. Their vision of the field is too remote from the one I wish to plow. So, I never had any illusion of being able to forge a union with them. Nonetheless, I contend that the field I am carving up (along with Frederique de Vignemont and with Gallese & Sinigaglia [2011]) is one that could well attract many members of the cognitive science community, in particular, practitioners of orthodox cognitive neuroscience. It should attract them, I believe, because it provides unity to an impressive array of empirically established phenomena with roots in bodily cognition. Yes, I would indeed hope to precipitate a unified vision of this subfield of cognitive science at some point down the road. But I do not expect this consensus to include all of the current proponents of EC, including all those represented by Shapiro’s list of ten examples. The orientation of roboticists and dynamic systems theorists, for example, is quite disparate from mine. It is unlikely that we would achieve a shared conception of our problems and assumptions. One need not agree with Shapiro, however, that his ten examples are ideally selected to identify the “essence” of EC. I could easily assemble quite a different ten-item list from those whose work I endorsed in the first two sections of this chapter. As Shapiro himself acknowledges, the history of science is replete with cases in which a scientific field (e.g., alchemy) displays shared assumptions at one point in time about the phenomena worthy of study, yet these assumptions are ultimately abandoned and replaced by others (e.g., those of modern chemistry). This scenario could easily be replicated in the area of EC.

Shapiro's discontent with my approach seems to be triggered partly by its moderation and partly by its brain-centrism. I see no reason to apologize for its moderation. (For many cognitive scientists, moreover, it is already too radical.) A first point to make about its brain-centrism is that such features emphasized by other embodiment theorists as nontrivial causal spread, situatedness, and embeddedness are also compatible with EC as I present it. I don't address these topics in the target article, but neither am I forced to reject them. (They might stand on their own without any linkage to my sense of embodiment.) Also worth mentioning—obvious though it is—is that, as far as anything I discuss goes, it is certainly possible (indeed, likely) that the *contents* of mental representation are very much a function of what they causally interact with. So envatted brain states would not have the same contents as brain states of ordinary embodied brains.

Shapiro's principal complaint, perhaps, is that my way of framing EC, with its adherence to brain-centrism, threatens to divest the embodiment movement of its most exciting and distinctive departure from orthodox cognitive science. Perhaps; but excitement isn't everything, especially in science. In any case, I think there is a pretty exciting and unorthodox theme in my story, which may be crystallized in the exaggerated slogan: "In the beginning, what we represent is our own body." Spelled out more cautiously, it's the thesis that a significant amount of human cognition has its origin in representations of one's own body.<sup>6</sup> This is not a trivial thesis. Indeed, most contemporary mainstream cognitive scientists probably reject it, and it certainly was not on the horizon for the first several decades of cognitive science. If it currently lacks shock value, this is because many relevant findings in cognitive science (and especially cognitive neuroscience) are already in the literature and comprise a firm empirical foundation on which I am trying to erect a sound theoretical edifice. The existence of those findings reduces the shock value of the thesis, but they also contribute mightily to its epistemic support, to its probability of being true. This firm empirical foundation is hardly a shortcoming. Sheer novelty is not the aim of science.

Moving forward in Shapiro's discussion, we come to his question of whether B-code use is indeed both necessary and sufficient for embodiment. Shapiro begins this phase of his discussion by disputing the necessity contention. However, he attributes to me the view that "cognition could not be embodied in anything but a human being" (this volume, Chapter 3). What leads Shapiro to interpret me in this fashion? He does not quote any passage from my target article to support this interpretation. True, the focus throughout the paper was human cognition, but this does not warrant the inference that I intended to restrict EC to the human species.<sup>7</sup>

Shapiro envisages a cognition-endowed creature that does not use dedicated B-formats. What about this class of creatures? he asks. Would they have embodied cognitions? I say that these creatures would not have embodied

cognitions. This follows directly from the proposed definition of embodied cognition. If a creature has no B-formats, it has no embodied cognitions. The divide between embodied and nonembodied cognition, as I seek to draw it, is in terms of B-formats. Of course, other theorists might wish to draw the line between embodied and nonembodied cognitions differently. They might want to draw a distinction, for example, between mental tokens that represent states of the body and mental tokens with no such representational content— ignoring entirely the codes, formats, or systems that are used. This is certainly a possible demarcation criterion; it just isn't as interesting a divide when it comes to the central scientific aim of cognitive science—namely, to understand the architecture of human cognition. Because, as the evidence I have assembled indicates, human cognition *does* involve the massive redeployment of cognition systems originally used for bodily representation but applied instead to a variety of tasks.

Having probed the issue of whether B-formats are *necessary* for embodiment, Shapiro next turns to whether they are *sufficient* for embodiment. In the very next sentence, however, he turns to a different question: whether embodiment shouldn't be treated as a graded category rather than a dichotomous one. This is surely a possibility, though not an easy one to manage, I suspect. Nonetheless, it is a matter we should definitely put on the agenda for future consideration. Different ways of measuring greater and lesser embodiment should be explored, to see how they might be used to broaden the approach.

Yet a different worry is expressed in Shapiro's worry that perhaps *all* cognition is embodied on the criterion I am proposing. Here we contemplate a very high degree of embodiment not for individual tokens of cognition but for the class of cognition tokens as a whole. Here again Shapiro surprises me with the following reaction to the prospect that *all* cognitions are embodied:

This consequence is surprisingly plausible, and yet, if genuine, would magnify suspicions that Goldman fails to articulate the most salient aspects of embodied cognition—aspects that distinguish embodied approaches to understanding the mind from classically cognitive ones. (this volume, Chapter 3)

I don't quite follow this suggestion. Exactly why should a “totalizing” upshot— *all* cognitions are embodied—magnify suspicions about my proposal? Maybe the thought is better articulated a few paragraphs later, where Shapiro says that massive (universal?) embodiment would threaten to *trivialize* the notion of embodiment.

Why would universal embodiment threaten to trivialize the embodiment notion? Admittedly, a definition of embodiment should leave *conceptual* room for both embodied and nonembodied cognitions. And that is clearly accomplished by my definition (both the one in Goldman [2012] and the one offered

above, BFC). But leaving conceptual room for both doesn't and shouldn't preclude the possibility that all cognitions are in fact embodied. An appropriate analogy here is the notion of the physical. A definition of "physical" should leave conceptual room for the possibility that both physical and nonphysical things exist. Still, it might turn out to be the case, as a factual matter, that all existing things are physical. Each of the putatively nonphysical things (God, numbers, thoughts, etc.) might turn out either (i) not to exist or (ii) to be physical after all. Why would this trivialize the notion of the physical?

Shapiro's final comment features a positive proposal to strengthen the proposed notion of EC. Instead of saying that it suffices for a cognition to qualify as embodied that it reuse a B-code, the test for embodiment (via reuse) should require both that the cognition reuse a B-code *and* that it reuse the B-code in question *because* this B-code has the function of representing certain bodily states.<sup>8</sup> This might indeed be a helpful addition. However, I suspect that the addition wouldn't make a big difference to the extension of cognitions that qualify as embodied, simply because the proposed extra condition would almost invariably be satisfied when the original condition of reuse is satisfied.

## Notes

1. A related thesis is the "extended mind" thesis, which construes the representing mind as including what is usually considered part of the environment. This approach is developed by Andy Clark and David Chalmers (Clark, 1997; Clark & Chalmers, 1998). However, I view this thesis as somewhat orthogonal to the embodied cognition thesis and won't try to work it into my discussion.
2. It is easy to find writers in the embodied cognition mold who, in the same work, vacillate between characterizations of their thesis in terms of the role of the body per se and some sort of *knowledge* or *representation* of bodily interrelationships. For example, Noe (2004) sometimes characterizes his "enactive" approach as the view that "perceiving [is] a way of acting" (2004: 1) and the view that "perception is a species of skillful bodily activity" (2004: 2). However, on the very same page of the latter quotation, he characterizes his view as holding that our ability to perceive "is constituted by . . . our possession of sensorimotor *knowledge*" (2004: 2, italics added). This formulation strongly implies that *representations* of sensorimotor relationships are what comprise perception.
3. That they are also activated in connection with language comprehension was a surprise, and the question of their function in this connection was the subject of the debate with Mahon and Caramazza considered previously. That this seems to be an *added* function is a topic to which we shall return.
4. Similar theories have been advanced in the literature, Anderson notes, by at least three other researchers or pairs of researchers: (1) Gallese and Lakoff's (2005) "neural exploitation" hypothesis; (2) Hurley's (2008) "shared circuits" model; and (3) Dehaene's (2005) "neuronal recycling" theory.
5. This is a somewhat improved formulation of the definition stated in Goldman (2012: 73).
6. This thesis is quite clearly presented by Antonio Damasio (2003). See especially chapter 3 on feelings.
7. I do want to avoid species chauvinism, as Shapiro suspects. But that is easily done with the recognition that other types of species, with different, non-neural material substrates, might have bodily formats and might reuse them for other purposes.

8. I give a tweak to Shapiro's proposal because I think it (slightly) enhances the appropriateness of the added condition.

## References

- Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence* 149: 91–130.
- Anderson, M. L. (2007). The massive redeployment hypothesis and the functional topography of the brain. *Philosophical Psychology* 21 (2): 143–174.
- Anderson, M. L. (2008). Circuit sharing and the implementation of intelligent systems. *Connection Science* 20 (4): 239–313.
- Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences* 33: 245–266.
- Barsalou, Lawrence W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 577–609.
- Barsalou, Lawrence W. (2008). Grounding symbolic operations in the brain's modal systems. In G. R. Semin and E. R. Smith, eds., *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches* (pp. 9–42). Cambridge: Cambridge University Press.
- Brooks, Rodney. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, MA: MIT Press.
- Casasanto, D., and Dykstra, K. (2010). Motor action emotional memory. *Cognition* 115 (1): 179–185.
- Clark, Andy. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58 (1): 7–19.
- Craig, A.D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews Neuroscience* 3: 655–666.
- Damasio, Antonio. (2003). *Looking for Spinoza: Joy, Sorrow and the Feeling Brain*. Orlando, FL: Harcourt.
- Dehaene, S. (2005). Evolution of human cortical circuits for reading and arithmetic: The “neuronal recycling” hypothesis. In S. Dehaene, J.-R. Duhamel, M.D. Hauser, and G. Rizzolatti, eds., *From Monkey Brain to Human Brain* (pp. 131–157). Cambridge, MA: MIT Press.
- de Vignemont, Frederique, and Jacob, Pierre. (2012). What is it like to feel another's pain? *Philosophy of Science* 79(2): 295–316.
- Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Gallese, Vittorio, and Lakoff, George. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology* 22 (3–4): 455–479.
- Gallese, Vittorio, and Sinigaglia, Corrado. (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences* 15 (11): 512–519.
- Glenberg, A.M., and Kaschak, M. F. (2002). Grounding language in action. *Psychonomic Bulletin and Review* 9: 558–565.
- Goldman, Alvin I. (2012). A moderate approach to embodied cognitive science. *The Review of Philosophy and Psychology* 3 (1): 71–88.
- Goldman, Alvin I., and de Vignemont, Frederique. (2009). Is social cognition embodied? *Trends in Cognitive Sciences* 13 (4): 154–159.
- Hauk, O., Johnsrude, I., and Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41: 301–307.
- Heidegger, Martin. (1962). *Being and Time*. New York: Harper Collins.
- Hilditch, D. (1995). At the heart of the world: Merleau-Ponty and the existential phenomenology of embodied and embedded intelligence in everyday coping. PhD thesis, Department of Philosophy, Washington University, St. Louis, MO.
- Hurley, S. L. (2008). The shared circuits model: How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences* 31 (1): 1–58.

- Lakoff, George. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lakoff, George, and Johnson, Mark. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
- Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology—Paris* 102: 59–76.
- Martin, A., Wiggs, G. L., Ungerleider, L. G., and Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature* 379: 649–652.
- Merleau-Ponty, Maurice. (2002). *The Phenomenology of Perception*. New York: Routledge.
- Noe, Alva. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- Prinz, Jesse. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press.
- Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience* 6: 576–582.
- Rizzolatti, G. and Sinigaglia, C. (2006). *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Oxford: Oxford University Press.
- Thelen, E., and Smith, L. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.