Embodied Functionalism and Inner Complexity: Simon’s 21st-Century Mind

Robert D. Rupert

University of Colorado at Boulder and University of Edinburgh

*I. Introduction*

One can hardly underestimate Herbert Simon’s influence on contemporary cognitive science and empirically oriented philosophy of mind. Working with collaborators at Carnegie Mellon and the Rand Corporation, he wrote Logic Theorist and General Problem Solver (GPS) and thereby helped to set the agenda for early work in Artificial Intelligence (AI) (Simon and Newell 1971). These projects also provided AI with some of its fundamental tools: in their work in the 1950s on Logic Theorist and other programs, Simon and colleagues invented list processing (which, in John McCarthy’s hands, became LISP – McCarthy 1960), while the conceptual framework of GPS gave birth to production systems (and became SOAR – Rosenbloom et al. 1991). In 1981, John Haugeland included Newell and Simon’s computationalist manifesto (“Computer science as empirical enquiry: Symbols and search” – Newell and Simon 1976) in his widely read anthology *Mind Design* (Haugeland 1981). As a result, the names of Newell and Simon became, in the philosophical world, nearly synonymous with the computational theory of the mind – at the top of the list with Fodor’s (1975) and Pylyshyn’s (1984). Moreover, in their early work, Simon and the Carnegie-Rand group emphasized the relative independence of an information-processing-based characterization of thought from the material components of the system so engaged (Newell, Shaw, and Simon 1958a, 51, Newell, Shaw, and Simon 1958b, 163, cf. Vera and Simon 1993, 9). Prominent philosophers, most notably Putnam (1960, 1963, 1967) and Fodor (1974), reified this distinctively functional level of description, thereby formulating the late-twentieth century’s dominant metaphysics of mind, functionalism, according to which mental states are, by their nature, multiply realizable functional states.

 The list of Simon’s influential ideas extends much further than this, however. Arguably, Simon’s work on hierarchical structure and the near-decomposability of complex systems (1996, Ch. 8) planted the seeds of modularity-based thinking, which blossomed in the work of Marr (1982), Fodor (1983), and evolutionary psychologists (Barkow, Cosmides, and Tooby 1992, Pinker 1997). In addition, Simon’s (1996) emphasis on satisficing and the bounded nature of rationality set the stage for the fascinating research programs of such varied figures Gerd Gigerenzer (2000), Christopher Cherniak (1986), and Ron McClamrock (1995). In fact, in conjunction with Kahneman and Tversky’s heuristics and biases program (Kahneman, Slovic, and Tversky 1982), Simon’s observations about the bounded nature of rationality yielded what is now the leading view of human cognition: it’s the work of a complex but nonoptimized system that manages surprisingly well by deploying its limited resources in a context-specific way – that is, in a way that exploits a grab-bag of relatively domain-specific shortcuts that are reasonably reliable given the environments in which they’re typically employed. At the same time, Simon sometimes emphasized adaptive rationality – the idea that, under a wide range of circumstances, intelligent human behavior is, given the subject’s goals, a straightforward function of the structure of the task at hand – a theme appearing in the work of such luminaries as Dan Dennett (1987) and John Anderson (1990). Simon also articulated a vision of intelligent behavior as the product of simple, internal mechanisms interacting with a complex external environment, and in doing so, inspired nearly two generations of philosophers and cognitive scientists who take intelligence to be the by-product of, or to emerge from, bodily interaction with the environment (Clark 1997, Brooks 1999). In addition, Simon (1996, 88, 94, 110) was perhaps the earliest working cognitive scientist explicitly to place aspects of internal processing on cognitive par with aspects of the environment external to the organism, a central theme in Andy Clark and David Chalmers’s enormously influential essay “The Extended Mind” (Clark and Chalmers 1998). There could scarcely be a more humbling list of one thinker’s achievements, and little has been said about Simon’s contributions to our understanding of search algorithms, economics, design, or management!

 As impressive as this list is, one might nevertheless wonder whether Simon’s views can be integrated into a single overarching vision of human cognition – of its structure, workings, and relation to the environment. Motivated by this kind of concern, I focus, in what follows, on an apparent inconsistency in Simon’s thinking, one that stands out especially clearly against the backdrop of decades of accumulated empirical results in what is sometimes known as ‘embodied cognitive science’ (Varela, Thompson, and Rosch 1991, Clark 1997, 2008b, Lakoff and Johnson 1999, Rowlands 1999, Gallagher 2005, Gibbs 2006, Rupert 2006, 2009, Shapiro 2010, Wilson and Foglia 2011). At first blush, embodied results seem to support certain strands of Simon’s thought at the expense of others. I argue, however, that even in cases in which Simon’s pronouncements about the mind were premature or his emphasis out of balance (e.g., not sufficiently oriented toward the body-based contributions to human cognition), many of his own views about the mind provide the necessary corrective and allow him to accommodate – even anticipate – embodiment-oriented insights. Simon wears an early, and perhaps slightly misshapen, version of the coat we should all gladly wear today, that of the embodied functionalist.

*II. The Tension*

*II.A Stage-setting and the dialectic*

Simon’s views about the human mind fall somewhat neatly, though not perfectly, into what might appear to be two mutually antagonistic clusters. In the first cluster, one finds views associated with what’s sometimes thought of as orthodox computationalism, according to which (a) the human mind operates in essentially the same way as does a human-engineered all-purpose digital computer and (b) we best understand the functioning of the human mind by focusing on computational properties and processes defined at a level of generality that subsumes both the human mind and the full variety of human-engineered all-purpose digital computers. In the other cluster sit views more sensitive to the details of the human condition, including views that seem to register limitations on the computer metaphor of the mind and direct the attention of cognitive scientists toward interactions among fine-grained (and thereby largely distinctive) aspects of human brain and extraneural body, as well as to the interactions of the human organism with the environment beyond its boundaries.

 In the remainder, I further articulate these clusters of ideas and attempt to neutralize the apparent tension between them, but first a few words about motivation. After all, why should we care whether Simon’s corpus presents a coherent picture of the human mind? What’s it to us, today? Readers who care deeply about the history of ideas have, by dint of that commitment alone, sufficient reason to reconstruct and evaluate Simon’s vision of the mind. In addition, there is a related issue of fairness: to the extent that Simon’s work set the stage for the development of widely held views in contemporary cognitive science – perhaps even embodiment-oriented views – Simon should be duly credited; more generally, we should want to acknowledge fully Simon’s contribution to the course of cognitive science.

 Perhaps of greatest interest to many readers, however, is the possibility that, by revisiting Simon’s views on mind and cognition, we clarify and help to resolve current debates in the philosophy of cognitive science. The present chapter aspires to this goal by examining relations between contemporary embodiment-oriented cognitive science and the theoretical commitments of historical cognitivism. Many contemporary embodiment theorists (Varela, Thompson, and Rosch 1991, van Gelder 1995, Glenberg 1997, Lakoff and Johnson 1999, Gibbs 2006, and various essays in Stewart, Gapenne, and Di Paolo 2010) claim that their view, qua embodied view, stands in stark contrast to the (purportedly waning) computational functionalist orthodoxy[[1]](#footnote-1) in cognitive science. I maintain that this way of casting the contemporary debate about embodiment rests on a fundamental misunderstanding of functionalism and of the historical contributions of cognitivists. A careful examination of Simon’s views bolsters this charge. It is true that, in some of his definitive-sounding, synoptic pronouncements about the mind, Simon seems to claim that it can be understood in complete ignorance of its bodily basis (e.g., 1996, 58–59, 73–74); in fact, he sometimes uses (although not entirely approvingly) the language of disembodiment (1996, 22, 83) to describe cognition. Furthermore, given Simon’s status a founding contributor to cognitive science, these pronouncements take on the glow of canon, appearing to sit at the heart of the computationalist orthodoxy in cognitive science and providing fair foil to those primarily concerned with bodily contributions to cognition. But, to narrate Simon’s story thusly would be a travesty; one need not read far from his remarks about disembodiment and the negligibility of the neural in order to find mitigating observations and commitments, some of which explicitly acknowledge the contribution of the contingent details of our bodily materials to the determination of the contingencies of our cognitive processing. The subtlety of Simon’s view of the mind should thus give us pause whenever we read that orthodox cognitive science must be overthrown if we are to accommodate the embodied perspective. For, if the views of Simon, an ur-computationalist at the center of the cognitive revolution, can accommodate, and in fact lay the theoretical groundwork for, embodied cognitive science, one should wonder whether the stated target of today’s embodiment revolutionaries, as well as the ways in which they situate their own position in the broader cognitive scientific scheme, mustn’t somehow be premised on a mistake.

 To be clear, various aspects of contemporary cognitive science, including some embodiment-related results, effectively challenge some ideas associated with historical cognitivism: anyone inclined to think that intelligence is what all and only universal Turing machines engage in should have flown a white flag long ago. All the same, if such untenable claims don’t fairly represent the historical views of the ur-practitioners of cognitivism (e.g., Simon), then to object to such views is not to take issue with the cognitivist tradition or to show that the tools and perspectives central to the cognitive revolution must be supplanted by the fruits of a new, embodied movement.

*II.B Simon and computational functionalism.*

Orthodox cognitive science is associated with (although perhaps not fully exhausted by – see note 1) the views that (a) intelligence has its basis in computational processing and (b) there is a distinctively functional level of description, explanation, or reality apposite to the scientific study of cognition or the philosophical understanding of its nature. Many of Simon’s programmatic statements about mind, cognition, and cognitive science seem to reflect these core elements of computational functionalism.

*II.B.1 Physical symbol systems*

Simon endorses the Physical Symbol Systems Hypothesis (PSS) about intelligence (Newell and Simon 1976, 116), which appears to manifest an uncompromising computationalism. According to the PSS, only physical symbol systems are intelligent, and whatever they do (that’s sufficiently complex or organized) is the exercise of intelligence. What is a physical symbol system?

A physical symbol system holds a set of entities, called symbols…[It] possesses a number of simple processes that operate upon symbol structures – processes that create, modify, copy, and destroy symbols…Symbols may also designate processes that the symbol system can interpret and execute. Hence the programs that govern the behavior of a symbol system can be stored…in the system’s own memory, and executed when activated. (1996, 22, and see *ibid.*, 19)

In a nutshell, intelligence is the activity of all-purpose, stored-program computers.

*II.B.2 Computational models of problem solving*

There is no doubt that Simon intends such claims to cover human intelligence in mind: as much as Simon’s description of physical symbol systems might put the reader in mind of artificial, computer processing, these operations “appear to be shared by the human central nervous system” (*ibid*, 19).

 Moreover, it was clear from early on (Newell, Shaw, and Simon 1958b) that Simon saw the computational modeling of human intelligence to be an explanation of how humans solve problems. Simon and colleagues were out to demystify human thought; by appealing to information processing models, they meant to “explain how human problem solving takes place: what processes are used, and what mechanisms perform these processes” (*ibid.*, 151). In describing their work on programs that, for example, prove logical theorems, play chess, and compose music, Simon and colleagues referred to their “success already achieved in synthesizing mechanisms that solve difficult problems in the same manner as humans” (1958a, 3).

*II.B.3 Functionalism*

Functionalism in philosophy of mind is, in a nutshell, the view that cognition is as cognition does, the idea that the nature of mental states and cognitive processes is determined by, and exhausted by, the role they play in causal networks, not a matter of their being constituted by some particular kind of materials. Simon unequivocally endorses a functionalist understanding of computing: “A computer is an organization of elementary functional components in which, to a high approximation, only the function performed by those components is relevant to the behavior of the whole system” (1996, 17–18). This theme continues over the pages that follow, with repeated comparisons to the human case: “For if it is the organization of components, and not physical properties, that largely determines behavior, and if computers are organized somewhat in the image of man, then the computer becomes an obvious device for exploring the consequences of alternative organizational assumptions for human behavior” (*ibid*., 21).

 In Simon’s hands, the description of the functional states at issue – at least at the levels of programming relevant to cognitive scientific enquiry – lines up with our common-sense way of describing the domains in question (they are semantically transparent, in the sense of Clark 1989, 2001). This is reflected partly in Simon’s use of subject protocols to inform his computational modeling of human cognition (Simon and Newell 1971, 150, 152, Newell, Shaw, and Simon 1958b, 156), not merely as a stream of verbal data that must be accounted for (by, for example, our best models of speech production), but as reports the contents of which provide a reasonably reliable guide to the processes operative in problem-solving (cf. Newell, Shaw, and Simon 1958a, 40 n1). This is worthy of note because, in the philosophical imagination, functionalism has often been understood as a claim about the kind of mental states one can specify in everyday language and the appearance and operation of which, in oneself, can be tracked by introspection. This has tended to obscure the live possibility that one can construct computational models of at least fragments of human cognition (a) that deal in representations of features or properties that have no everyday expression in natural language and to which we have no conscious access and (b) that may well tell the entire story about the fragment of cognition in question. In other words, Simon’s work with protocols and his account of, for instance, the processes of means-end reasoning and theorem-proving – accounts that seem to involve representations that match naturally with the fruits of introspection – reinforce the image of Simon as a coarse-grained computational functionalist who makes no room for subtle, subconscious body-based processing. (Cf. Simon’s discussion of memory for chess positions [1996, 72–73]; Simon generates his description of the relevant cognitive process by auto-report and couches his account of the process in everyday chess-playing terms.) In contrast, many embodiment oriented models deal in features and processes difficult to express in natural language and better captured only mathematically or in a computational formalism that has no natural, everyday expression (cf. Clark’s [1989] characterization of connectionist networks as detecting and processing microfeatures). A computational functionalism of the latter, fine-grained sort is computational functionalism nevertheless.

*II.B.4 Multiple realization and metaphysical autonomy*

As a philosophical theory of the nature of mental states, functionalism entails multiple realizability (MR), that the very same mental or cognitive properties or states can appear in, be implemented by, or take the form of significantly different physical structures (different with respect to how the physical sciences would characterize those structures and their properties). Although Simon’s comments often seem oriented toward methods of investigation or kinds of explanation, he appears to be committed to metaphysical functionalism and the associated MR thesis as well.

 Newell, Shaw, and Simon express the view in this way:

We do not believe that this functional equivalence between brains and computers implies any structural equivalence at a more minute anatomic level (for example, equivalence of neurons with circuits). Discovering what neural mechanisms realize these information processing functions in the human brain is a task for another level of theory construction. Our theory is a theory of the information processes involved in problem solving and not a theory of neuronal or electronic mechanisms for information processing. (1964, 352 – quoted at Gardner 1987, 148; cf. Newell, Shaw, and Simon 1958a, 51)

And, from a more recent paper by Vera and Simon (a paper about situated cognition, no less!): “And, in any event, their physical nature is irrelevant to their role in behavior. The way in which symbols are represented in the brain is not known; presumably, they are patterns of neuronal arrangements of some kind” (1993, 9).

 Moreover, when Simon discusses the functional equivalence of computer and human, he says that “both computer and brain, when engaged in thought, are adaptive systems, seeking to mold themselves to the shape of the task environment” (1996, 83). Taking such remarks at face value, Simon seems to be talking about the very things and properties in the world, as they are, not merely, for example, how they are best described for some practical purpose.

*II.B.5 MR and methodological autonomy*

At the same time, we should recognize that Simon is often interested in matters methodological or epistemological, in how one should go about investigating intelligent systems or explaining the output of intelligent systems. He frequently talks about our interests, and about discovery, theory construction, and description. Even the equivalence referred to above could just as well be equivalence vis-à-vis our epistemic interests rather than equivalence in what exists independently of those interests.

 Frequently these metaphysical and epistemic messages run parallel to each other (perhaps because Simon has both in mind, and the metaphysical nature of the properties and patterns in question determines the important epistemic and methodological facts). For instance, he says that

many of the phenomena of visualization do not depend in any detailed way [note the use of ‘detailed’] upon underlying neurology but can be explained and predicted on the basis of quite general and abstract features of the organization of memory. (1996, 73–74)

Although one might wonder about the nature of the dependence being discussed, the passage’s emphasis on the predictive and explanatory value of postulated features lends itself naturally to a primarily epistemological reading.

 This is not an exercise in hair splitting. A failure to distinguish between the metaphysical and epistemic or methodological dimensions of computational functionalism can easily obscure the relation between computational functionalism as a theory of mind and computational functionalism as it’s sometimes been practiced, the former of which makes plenty of room for embodied cognitive science, even if some computational functionalists have not, in practice, focused on bodily contributions to cognition.

*II.B.4 Adaptive rationality*

The idea of adaptive rationality plays a central role in *Sciences of the Artificial*. After describing what they take to be the small number of parameters constraining human thought – including, for example, that human short-term memory can hold approximately seven chunks – Newell and Simon say:

[T]he system might be almost anything as long as it meets these few structural and parametral specifications. The detail is elusive because the system is adaptive. For a system to be adaptive means that it is capable of grappling with whatever task environment confronts it…its behavior is determined by the demands of that task environment rather than by its own internal characteristics (1971, 149; and see Simon 1996, 11–12)

On this view, we can, for many purposes, think of the internal workings of the cognitive systems as black boxes; but for a small number of parameters determined by their material constitution, human cognitive systems – in fact, any intelligent systems – are organized in some way or another so as to adapt to task environments. As intelligent systems they thus exhibit similar behavior across a wide range of circumstances (where goals are shared).

*II.B.5 Simplicity of the inner*

Simon opens a central chapter of *Sciences of the Artificial* with the tale of an ant picking its way across a windblown beach toward its nest, encountering many small obstacles along the way. To an onlooker, the ant’s path might seem to reflect a complex internal process. According to Simon, though, the path’s “complexity is really a complexity in the surface of the beach, not a complexity in the ant” (1996, 51). He extends the moral of the story to human cognition and behavior: “Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves” (1996, 53).

 This is a striking claim, not entailed by adaptivity; for, various highly organized systems might all nevertheless be organized so as to behave adaptively, and one might even think it necessary that a system have a complex structure in order to respond effectively in a wide range of task environments. I emphasize the claim of inner simplicity at this point in the discussion because the claim seems to stand so deeply at odds with the embodied program, which takes human cognition, in all of its nuance and complexity, to be largely a function of fine-grained facts about the distinctive structural and causal organization of the inner workings of the human system.

Throughout *II.B*, I have portrayed Simon as a representative of a brand of computational functionalism often derided by embodied theorists for its utter disregard for the material basis of cognition and for its tendency instead to fetishize such projects as a formal analysis of task domains and the exploration of algorithms of search through them, algorithms that have no connection to distinctively human, body-based strategies.

*II.C Embodied functionalism and contingency-making*

The preceding portrayal significantly misrepresents Simon’s approach to the mind, and here’s a way to begin to see why. One of the most important aspects of the embodied view is a commitment to what might be reasonably called ‘contingency making’. On the embodied view, contingent facts about our bodily existence color and shape human cognition (and presumably, something similar holds for other creatures as well). Humans have a particular body shape and orientation and particular networks of muscles that interact in a finely-tuned orchestra of counterbalancing forces to keep the organism alive and functioning in the face of ongoing perturbation by environmental forces. Moreover, patterns of neural firings involved in such orchestration provide resources to be co-opted for other cognitive purposes, from the encoding of items in working memory (Wilson 2001) to the metaphorical understanding of abstract concepts (Lakoff and Johnson 1999). On this picture, the resources used in real time by human organisms – in all of their *ad hoc*, makeshift, exapted, kludgey, but stunningly effective, glory – bear the marks of this distinctive body-based orchestra. Thus, according to embodiment theorists, there is simply no way to understand what’s going on with human cognition absent the thorough investigation of these contingency-making, bodily contributions.

 Simon’s picture, too, is rife with contingency making, arguably of the bodily sort (although he does not often use such language), which places him – or at least the second cluster of his views – squarely in a camp with the embodiment theorists. In fact, one might reasonably contend that Simon’s uncompromising interest in the contingencies of the human case helped to set the stage for the embodied movement.

 It is the burden of *Section III* to make a detailed case for this claim. But, first, a door must be pried open. Don’t embodiment theorists routinely criticize, reject, and even demonize computationalism, functionalism, and multiple realizability? How could I possibly, in all seriousness, present Simon as a proto-embodiment theorist, given his clear commitments to computational functionalism and related views, as documented above? The answer requires that we revisit these much-derided orthodox positions and clarify their relation to embodied cognitive science. In doing so, we clear the way for a proper understanding of Simon’s view and of embodied functionalism, generally speaking.

*III. Fine-grained Functionalism: Respecting the Bodily Source of Functionally Specified Contingencies*

Many embodiment theorists criticize functionalism and computationalism as “disembodied” and reject them outright. If these critics are correct, then, Simon – *qua* ur-computationalist committed to functionalism, computationalism, multiple-realizability, and the irrelevance of the physical – can’t possibly be cast as an embodiment theorist, not without radical revision or wholesale denial of significant aspects of his thought.
 This way of framing the dialectic gains no traction against the conciliatory picture of Simon’s thought being constructed here, for it rests on a misunderstanding of computationalism and functionalism (cf. Rupert 2006, 2009 ch. 11, Clark 2008a, Shapiro 2010). What, after all, does it mean to claim that functionalist and computationalist approaches are disembodied? One version of the complaint seems to be that the physical material of the body or brain doesn’t play a significant metaphysical role in the functionalist’s or computationalist’s theory of mind or cognition; the body has been left out. If that, however, correctly represents the operative complaint, embodied theorists have misunderstood the metaphysical relation of their views to classical cognitive research. According to the classic versions of functionalism, the physical body (the brain, in particular, but this applies to whatever matter realizes the computational system) determines, in the strictest metaphysical sense, which functions the human cognitive system computes. That is the very nature of the realization-relation. Thus, according to classical computational functionalism, the body is the root and determinant of our cognitive being. Thus, unless we construe embodied theories as type-type identity theories – which raises a host of its own problems (Clark 2008a) – embodied views are in precisely the same boat metaphysically as functionalism. Embodied views in cognitive science are, in fact, functionalist views. They are, however, fine-grained functionalist views, in that they tend to build fine details of the workings of the human brain, body, and world into their functionalist models, by individuating the functional properties in those cognitive models in a way that reflects the details of the causal profiles of the relevant physical materials (although stopping well short of any attempt to capture the full causal profiles of those materials). So long as more than one collection of physical matter could exhibit the same, relevant causal profile as described in an embodied model of cognition, the processes in question fit functionalist metaphysics perfectly. They may not be recognized by commonsense or analytic functionalists, but they are functionalist in the sense relevant to cognitive science; they are the fruits of a psychofunctionalist approach (Block 1978, 269), which looks to science – embodied cognitive science, as it turns out – to characterize the functional properties in question.

 The embodied approach recommends a different methodology than was pursued by many classical computationally oriented cognitive scientists. The embodied functionalist lets such things as bodily activity be her guide, epistemically, when attempting to figure out, for instance, which algorithms (or other abstract processes – a nod to those who lean toward dynamical-systems-based modeling [e.g., Chemero 2009]) govern human cognition. This has no bearing on the truth of functionalism, however (except to confirm functionalism by the success of the embodied strategy!), but rather it stands in opposition to a certain empirical bet that many computationalist-functionalists made in the early days of cognitive science: that the relevant algorithms and the location of the machinery that executes them could be identified from the armchair or by reflection on the way everyday people typically talk and think about human thought processes. (Dennett makes kind of point about neuromodulators and neuroanatomy in the context of a discussion of functionalism and consciousness [2005, 17–20].)

 A now relatively common way to capture the differences among functionalist views is to recognize a continuum running from coarse-grained to fine-grained versions of functionalism. A given functionally individuated state can be individuated by a simple pattern of relations among very few states, inputs, and outputs (consider the stock example of Coke machine – Block 1978, 267) to the massively populated space of states, including millions of possible input and output states (as is the case with the typical human cognitive state the presence of which is determined by the incredibly complex network of functional and computational states realized by and thus determined by the brain). In his official pronouncements, Simon seems to commit himself to a coarse-grained functionalist view, and it’s clear enough why. If human internal cognitive operations are so few and their workings straightforward, constrained only by a small number of biologically set parameters easily measured and modeled using behavioral data alone, then human cognitive operations are broadly multiply realizable and largely independent of the fine details of neural structure and processing. But, Simon could be wrong about this even if functionalism and computationalism are true. Moreover, the turn toward a fine-grained computational functionalism – even one that focuses on body-related contributions – may be inspired by some of Simon’s own work.

*IV. The Body of Simon’s Work*

In this section, I explore Simon’s views that fall into the second of the two clusters identified at the outset. Simon’s explicit methodological pronouncements often jibe with what one would expect from an embodiment theorist, and many of the theoretical constructs central to his work can be most fruitfully understood as ways of discovering the body’s contingent effects on human cognitive processing. Or, so I argue in the remainder.

*IV.A Simon, environmental interaction, and bodily determined conceptual variation*

What methodology does an embodied perspective in cognitive science call for? One common embodied approach tracks the cognitive effects of interaction with the environment, either on an evolutionary, learning, or developmental scale (Rupert 1998); this can be done via simulation of embodied agents (Beer 2003, Husbands, Harvey, and Cliff 1995), construction of robots (Brooks 1999, Steels and Spranger 2008), or observation of human infants (Thelen and Smith 1994). Much of this work focuses not only on the diachronic issue of changes in systems over time, but also on the real-time interaction with the environment, seeking to uncover the contribution that bodily motion and interaction with the environment make to intelligent behavior (Kirsh and Maglio 1994). Simon’s remarks about the ant would appear to provide clear inspiration for this kind of interactive work, which is central to the history of embodied cognitive science.

 Of significant interest, too, is the modeling of fine-grained details of the neural contributions to bodily control (Grush 2004), including the contribution of neural representations (or other mechanisms or processes) that track bodily interaction with the environment (Iriki, Tanaka, and Iwamura 1996). Such a project proceeds by combining at least two forms of reasoning: (a) the use of behavioral data together with our best guesses, given the task, about which functions are likely being performed by areas of the brain active during task performance and (b) sensitivity to activity at the neural level during task performance to help us generate and winnow among theories of the functions being carried out, which might be quite fine-grained, species-specific, and based on bodily contingencies (while still being multiply realizable, at least in principle).

 Simon and Newell state this strategy explicitly as part of their 11-point strategy for explaining human problem solving:

Begin to search for the neurophysiological counterparts of the elementary information processes that are postulated in the theories. Use neurophysiological evidence to improve the problem-solving theories, and inferences from the problem-solving theories as clues for the neurophysiological investigations. (1971, 146)

Thus, there is nothing contradictory in Simon’s (and Newell’s) being a proto-embodied-functionalist – with regard to method, not only metaphysics – while at the same time pressing ahead with a largely information-processing-based research program (*ibid.*, 157–158), one inspired by behavioral data and *a priori* theorizing about the functional organization of the system producing that behavior.

 Note, too, their endorsement of contingency:

These properties-—serial processing, small short-term memory, infinite long-term memory with fast retrieval but slow storage—impose strong constraints on the ways in which the system can seek solutions to problems in larger problem spaces. A system not sharing these properties—a parallel system, say, or one capable of storing symbols in long-term memory in milliseconds instead of seconds—might seek problem solutions in quite different ways from the system we are considering. (*ibid.*, 149)

“Different bodies, different concepts” is one of the most widely known themes from the embodiment literature (Shapiro 2010), and according to leading embodiment theorists (Lakoff and Johnson 1999), processing profiles individuate concepts (that is, determine whether two concepts are the same or different). Thus, Simon and Newell seem to be marching in step with the embodiment theorists: a system with a significantly different material basis might well solve problems in different ways, which entails difference in processing; and, since the concepts at issue are individuated by their processing profiles, solving problems in different ways would seem to entail the use of different concepts.

 Although Simon sometimes seems to advocate for the “disembodiment of mind” (Simon 1996, 83), Simon explicitly qualifies such remarks: “It would be unfortunate if this conclusion were altered to read that neurophysiology has nothing to contribute to the explanation of human behavior. That would of course be a ridiculous doctrine…It is to physiology that we must turn for an explanation of the limits of adaptation” (*ibid.*). The human system doesn’t adapt perfectly to the problems it faces, and when its behavior is not driven optimally by the structure of the problem at hand, it is because of the constraints placed on it by neurophysiology.

 Simon’s two recurring examples of such constraints are that working memory can hold only about seven chunks of information and that it takes about eight seconds for an item to be transferred into long-term memory. Simon should have known better, however, given his various commitments and insights, than to emphasize the smallness of the number of such constraints and to push them to the theoretical margins, as merely the limits of adaptation. He acknowledges, for example, that “Neurophysiology is the study of the inner environment” (*ibid.*). And although he seems to identify explicitly only two or three interesting parameters determined by that inner environment – in his official pronouncements about said number, at least – he seems to invoke many such parameters and identify many effects of such parameters across the range of his work. Moreover, his general theoretical outlook gives him permission to identify many more such parameters – as many as the data demand – as is reflected by his actual practice.

*IV.B The body ascendant*

Beyond the general themes discussed in *IV.A*, there are various ways in which Simon’s own theorizing seems to encourage an embodied perspective, as this subsection demonstrates.

*IV.B.1*. *Stress, duress, and taxing environments*

Notice the way in which Simon and Newell talk about the discovery of parameters: “Only when the environment stresses its capacities along some dimension…do we discover what those capabilities and limits are, and are we able to measure some of their parameters (Simon, 1969, Ch. 1 and 2).” (1971, 149) And, unsurprisingly, their flagging of Simon 1969 (which I’ve been citing in its most recent edition, 1996) is spot on. When a system is taxed, Simon claims, or makes a mistake, the properties of the system’s material substrate, which would not otherwise show themselves, become behaviorally manifest (1996, 12, 58, 83).

 Interestingly, contemporary embodiment theorists often emphasize that cognition is time pressured and evolution kludge and that this affects cognition through and through. What embodiment theorists add to Simon’s perspective, then – which addition Simon can naturally accommodate – is that the limiting properties of the inner system virtually always show through. Because thought is commonly time pressured – or in some other way pressured by context, including internal context – the human cognitive system is almost always being taxed, and thus significant and contingent aspects of the material substrate of cognition are continually on display. A constant stream of external stimuli together with the ongoing fluctuations in internal context (neuromodulators, hormone levels, neural correlates of emotional states and moods) create a shifting cognitive context, a series of “perturbing influences” on cognitive processing; and under these conditions, the fine-grained physical properties of our brains – and bodies and external environments, perhaps – reveal themselves in all sorts of ways relevant to our understanding of human performance and intelligent behavior.

*IV.B.2*. *Beyond duress*

I’ve tried to extend the scope of Simon’s category of a system under duress, so that it covers a wide range of cases in which embodied contributions to cognition reveal themselves. Circumstances involving what is more clearly conceived as breakdown also reveal fascinating aspects of the human cognition, in ways that Simon would have no reason to resist. Over the decades, cognitive scientists have learned no small amount about the human cognitive system – its architecture and component mechanisms – from the study of subjects suffering from autism spectrum disorders, Capgras syndrome, hemispatial neglect, prosopagnosia, Balint’s syndrome, and many more disorders.

 Experimental manipulations, too, have led to the discovery of a variety of parameters (and their values) as well as systemic properties that emerge from interaction among sometimes quirkily functioning component parts, all of which Simon can gladly take on board. Consider, for instance, Pylyshyn’s proposal that there are four FINST pointers (Pylyshyn 2000), or studies showing that tactile discrimination becomes more sensitive when experimenters activate a visual representation of the area being touched (Serino and Haggard 2010), or that one’s report of the timing of one’s conscious intention to press a button shifts when pre-SMC is subject to TMS *after* one has pressed the button (Lau, Rogers, and Passingham 2007). Presumably, the material basis of human cognition, including materials that have more or less to do with specifically sensory or motor contributions to cognition, determine the presence in humans of such cognitively relevant architectural elements and their specific causal profiles.

 In his attempt to understand the complexities of human problem solving, Simon himself identifies numerous contingent architectural elements of the human cognitive system and phenomena that would seem to arise contingently from the interaction of these elements, and many of these seem amenable to an embodied treatment. Among Simon’s most enduring contributions to cognitive scientific theorizing are his notions of bounded cognition and the associated idea of satisficing. Humans have limited cognitive resources, limited long- and short-term memory capacities, for example. But, by satisficing (looking for a satisfactory, or “good enough,” solution), we compensate for the limited (that is, bounded) nature of our cognitive systems and thereby solve (well enough) problems we couldn’t easily (or even possibly!) solve optimally, by, for example, exhaustive search through, and evaluation of, all possible options. Most of us don’t undertake the project of finding the very best wine to go with our dinner; instead, we find one that’s good enough. We achieve this partly by employing search heuristics (we buy a wine we’ve heard of, that seems priced about right, and that has a high Wine-Spectator score posted beneath it).

 At this point, one might reasonably wonder whether Simon wasn’t onto one of the most substantive insights of the embodiment-based literature. Think of how deeply nonoptimizing is the human use of heuristics and biases. Cognitive scientists with time on their hands and lots of computing power can explain to us why the shortcuts we use represent reasonable approaches to time- and resource-pressured problem solving. But, *we*, in situ, can’t do such calculations; we would lose the advantage of employing heuristics to satisfice if we were to attempt calculate the costs and benefits of using those heuristics in a given case (McClamrock 1995). Instead their use must come naturally, in some sense. And, the only plausible ways for them to come naturally are the ones emphasized by embodiment theorists: we solve problems by relying *automatically* on shortcuts built into the neural system, into our bodily structure, and into our environmental tools, and by learning to exploit ways in which the fine-grained details of our bodily and neural structures interact with fine-grained aspects of our environments.

*IV.B.3*. *More of Simon’s own*

Many more of Simon’s theoretical constructs or posits dovetail with, or even lend support to, an embodied perspective:

a. Heuristics and search. In a discussion of search strategies in chess, Simon and Newell say, “The progressive deepening strategy is not imposed on the player by the structure of the chess task environment. Indeed, one can show that a different organization would permit more efficient search” (1971, 153). And note that, as their following paragraphs make clear, Simon and Newell are interested, not only in the way humans happen to play chess, but also the way in which the memory resources available – themselves bodily determined – help to determine the search strategy humans choose.

b. Along related lines, Newell and Simon identify in subjects a tendency, as part of a planning phase, to abstract from the details of a particular problem, without (somewhat surprisingly) maintaining a clear distinction between the abstracted space used in planning (or strategizing) and the problem space ultimately appropriate to the problem at hand (1971, 156). An embodied theorist might wonder whether the human moves easily between such spaces because, in some sense (geometrically, perhaps?), the bodily resources for one way of thinking are nested within the resources for the other. And, other things being equal, there’s no reason Simon shouldn’t pursue this kind of answer to his questions about the selection of problem spaces to work in – perhaps seeing ease of migration from one space to a more abstract version of it as another body-determined parameter.

c. The human’s retrieval of information from long-term memory is itself governed by further processes of interest (Simon 1996, 82). This process includes the use of associative memory, which, according to Simon, is best understood as the manipulation of list structures. This central aspect of the organization of the human mind, which shows its distinctive character in experimental settings is plausibly a function of the fine-grained properties of the material basis of the human cognitive system. Moreover, it’s particularly plausible that the resources for such associations include images and motor patterns – the sorts of resources emphasized by embodiment theorists.

d. Simon observes and attempts to explain deviations from economic rationality: “Affected by their organizational identifications, members frequently pursue organizational goals at the expense of their own interests” (1996, 44). What drives such identification? How do analytically capable individuals become company men and company woman even though the individual irrationality of doing so is demonstrable? Presumably, this results from ways in which the material basis of human cognition contingently determines our affective states and processing, the contribution of which to cognition has come to be understood better and better in recent decades (see the literature on the Iowa Gambling task, for example – Bechara et al. 2005) and is often associated with the embodied perspective.

e. In Simon’s discussion of design (1996, 128–130), he emphasizes the importance of style, of the choices architects and composers face at various points in what are essentially processes of creative problem-solving. The results of the choices made – in effect, the expressions of different styles – can be seen as a contribution to generate-and-test cycles, according to Simon. The choices contribute to the generation of a variety of possibilities, the availability of which will, among its other benefits, increase the probability of access to a satisfactory option. And in the paragraph that follows, Simon extends this vision to the design of cities and economies.

 What, however, determines such differences in style? Learning history? Genetic differences? Perhaps, but how would such variations in history and genes create variation? Particularly when it comes to such intensely visual and spatial domains as design, one would think they do so by affecting the sorts of cognitive resources – sensori-motor routines, for example – of interest to embodied theorists.

f. Simon’s work suggests many more possibilities for the application of an embodied perspective: Why does the human production system employ the difference-reduction operators it does (1996, 94) rather than others? Why does human attention work in the way it does (*ibid.*, 143–144)? What is the bodily basis of our tendency to discount the future, which Simon himself recognizes as another “of the constraints on adaptation belonging to the inner environment” (*ibid.*, 157)?

 As in the cases discussed above – of associative memory, use of search heuristics, the contribution of varying styles to design processes, and the contribution of “irrational” factors to economic cognition – the profile of human processing seems both unaccounted for by adaptive rationality alone, partly because it is shot through with contingency. And, while these contingencies can be functionally and computationally (or at least mathematically) characterized, that hardly makes them elements of a universal disembodied rationality. The contingencies themselves are determined by the fine-grained contingencies of the bodily materials that realize the functions in question.

*IV.B.4*. *The co-opting of problem spaces*

According to Simon, “Every problem-solving effort must begin with creating a representation for the problem—a problem space in which the search for the solution can take place” (1996, 108). A subject’s choice of a problem space is way the “particular subject represents the task in order to work on it” (Simon and Newell 1971, 151). About these matters, Simon is characteristically honest, “The process of discovering new representations is a major missing link in our theories of thinking…” (Simon 1996, 109; cf. Simon and Newell 1971, 154).

 To my mind, questions about the acquisition and selection of problem spaces are some of the most fascinating in cognitive science. One such question is under what conditions the subject co-opts an already acquired representation of a problem space for use in the solution of a new problem. Simon discusses this issue in connection with a simple card game. After describing its rules, Simon notes that the game has essentially the same structure as the familiar game of tic-tac-toe (1996, 131–132), which, I imagine, comes as a surprise to most readers. But, once one has seen the structural equivalence, one can immediately transfer strategies for playing tic-tac-toe to Simon’s card game.

 Three comments are in order, two of which concern the connection to embodied theorizing and the ways in which embodied theorizing might help to answer some of Simon’s open questions about problem solving. First, many of Simon’s examples of problem-solving involve mathematical or otherwise symbolic reasoning. It’s worth noting, then, relatively recent work on the role of gesture in the development of problem-solving skills in arithmetic (Goldin-Meadow 2003). Second, questions to do with encoding – about representing the problem within the problem space – have also been given an embodied treatment. This research demonstrates the great extent to which symbolic reasoning depends on spatial arrangements in the external world and the way they interact with our perceptuo-motor resources (Landy, Allen, and Zednik 2014).

 My third observation distances the discussion somewhat from specific claims about mathematical or symbolic cognition. Rather it is to note the depth and scope of Simon’s observations regarding the co-opting of problem spaces. On Simon’s view, cognition is bounded, and, as a result, we take cognitive shortcuts. We do so partly by mapping new problems into a stock of existing problem spaces that are themselves amenable to the use of effective heuristics. This mapping of problems – typically more complex problems – into less complex or more familiar problem spaces is one of the primary themes of Daniel Kahneman’s recent overview of his life’s work (Kahneman 2011). This is also one of the central themes of the work of one of the champions of embodiment, George Lakoff. Much of Lakoff’s work has been an effort to show that a relatively small number of body-based source domains provide the materials for a metaphorical understanding of what we think of as abstract concepts and domains (Lakoff and Johnson 1999). It’s plausible enough, then, that important aspects of the availability of Simon’s heuristics, as well as their structural natures, are determined by the details of our bodily experience. And given the centrality of these issues to Simon’s picture of the mind, embodiment would emerge as central to a theory of cognition wedding, in this way, the work of Simon, Kahneman, and Lakoff.

 In the end, Simon’s vision need have very little added to it to become a deeply embodied position. And what must be added conflicts in no way with his big-picture theoretical commitments, qua computational functionalism, and is at least suggested by many of his other most important theoretical commitments.

*V. Conclusion*

In this chapter, I identified two clusters of Simon’s views, some which clearly align him computational functionalism and the others of which either implicitly or explicitly presuppose or support significant aspects of the embodied perspective. I also argued that the former views do not conflict with the embodied perspective. On balance, Simon’s picture of the mind emerges as a coherent precursor to contemporary embodied functionalism.

 There are broader lessons to be learned here as well. Virtually everyone in the philosophy of cognitive science, from orthodox computationalist to embodiment theorist, is in the same boat: they are all metaphysically disembodied[[2]](#footnote-2) and epistemically embodied, and always have been. Unless embodiment theorists choose to embrace type-type-type identity theory as a metaphysics of mind, their view is every bit as functionalist and disembodied as was early cognitivism. According to both groups, the physical materials fully determine what happens at the cognitive level but cannot be literally identified with those cognitive processes.[[3]](#footnote-3) Even those who reject computationalism as a modeling methodology (in favor of, say, of dynamical-systems-based modeling – Chemero 2009) do not identify their theoretical types with physical types: there is no brain state that is, literally, the one and only kind of state that is identical to, for example, a periodic attractor or an attractor well; all parties are up to their necks in multiply realizable kinds, at the very least in the sense that they make explanatory use of abstract kinds or properties that take a variety of forms (see Weiskopf 2004 for a related problem).

 At the same time, from the epistemic or methodological standpoint, cognitive science has always been concerned with embodiment. From Hebb (1949) to Lettvin, Maturana, McCulloch, and Pitts (1959) to Hubel and Wiesel (1962) to Sperry, Gazzaniga, and Bogen (1969) to Bliss and [Lømo](http://en.wikipedia.org/wiki/Per_Andersen) (1973) to Warrington (1975), the early decades of cognitive science saw a steady stream of highly influential work in neuroscience that was of broad interest specifically because it was taken to have cognitive implications. The real history of cognitive science – not the “disembodied” caricature of it – was of a piece with embodied functionalism, even though some major figures, including Simon in some moods, may have hoped they could generate computational models of cognition without attending to the bodily materials that they quite well knew determined those computational functions in humans.

References

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Lawrence Erlbaum.

Barkow, J., Cosmides, L., and Tooby, J. (eds.) (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (2005). The Iowa Gambling

Task and the somatic marker hypothesis: Some questions and answers. *Trends*

*in Cognitive Sciences* 9, 4, 159–62.

Beer, R. (2003). The dynamics of active categorical perception in an evolved model

agent. *Adaptive Behavior* 11, 209–43.

Bliss, T. V. P., and [Lømo](http://en.wikipedia.org/wiki/Per_Andersen), T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, 232, 331–356.

Block, N. (1978). Troubles with functionalism. In C. Savage (ed.), *Perception and cognition. Issues in the foundations of psychology, Minnesota studies in the philosophy of Science*, vol. 9 (Minneapolis: University of Minnesota Press), pp. 261–325.

Brooks, R. (1999). *Cambrian intelligence: The early history of the new AI*. Cambridge,

MA: MIT Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.

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

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

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science.* Oxford: Oxford University Press.

Clark, A. (2008a). Pressing the flesh: A tension in the study of the embodied,

embedded mind. *Philosophy and Phenomenological Research* 76, 1, 37–59.

Clark, A. (2008b). *Supersizing the mind: Embodiment, action, and cognitive extension*.

Oxford: Oxford University Press.

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

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT Press.

Fodor, J. A. (1974). Special sciences. *Synthese*, 28, 77–115.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Gallagher, S. (2005). *How the body shapes the mind.* New York: Oxford University

Press.

Gardner, H. (1987). *The mind’s new science: A history of the cognitive revolution.* New York: Basic Books

Gibbs, R. (2006). *Embodiment and cognitive science*. Cambridge: Cambridge University Press.

Gigerenzer, G. (2000). *Adaptive thinking*. Oxford: Oxford University Press.

Glenberg, A. (1997). What memory is for. *Behavioral and Brain Sciences* 20, 1–19.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think.* Cambridge: Belknap Press/Harvard University Press.

Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and

perception. *Behavioral and Brain Sciences,* 27, 377–396.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42, 335–46.

Haugeland, J. (ed.) (1981). *Mind design*. Cambridge, MA: Bradford/MIT Press.

Hebb, D. (1949). *The organization of behavior*. New York: Wiley & Sons.

Hubel, D. H., and Wiesel, T. N. (1962). Receptive Fields, binocular interaction and functional architecture in the cat’s visual cortex. *Journal of Physiology*, 160, 106–154.

Husbands, P., Harvey, I., and Cliff, D. (1995). Circle in the round: State space attractors for evolved sighted robots. *Robotics and Autonomous Systems* 15, 83–106.

Iriki, A., Tanaka, M., and Iwamura, Y. (1996). Coding of modified body schema

during tool use by macaque postcentral neurones. *Neuroreport*, 7, 2325–2330.

Kahneman, Daniel. (2011) *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., Slovic, P., and Tversky, A. (eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kirsh, D., and Maglio, P. (1994). On distinguishing epistemic from pragmatic action.

*Cognitive Science*, 18, 513–549.

Lakoff, G. and Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* New York: Basic Books.

Landy, D., Allen, C., and Zednik, C. (2014). A perceptual account of symbolic reasoning. *Frontiers in Psychology*, 5, 1–10 (Article 275).

Lau, H. C., Rogers, R. D., and Passingham, R. E. (2007). Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience*, 19, 1, 1–10.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., and Pitts, W. A. (1959). What the frog’s eye tells the frog’s brain. *Proceedings of the Institute of Radio Engineers*, 47, 11, 1940–1951.

Marr, D. (1982). *Vision*. New York, NY: W. H. Freeman.

McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, part I. *CACM,* 3, 4, 184–195.

McClamrock, R. (1995). *Existential cognition: Computational minds in the world*. Chicago: University of Chicago Press.

Newell, A., Shaw, J. C., and Simon, H. A. (1958a). The process of creative thinking. Rand Corporation technical report P-1320.

Newell, A., Shaw, J. C., and Simon, H. A. (1958b). Elements of a theory of human problem solving. *Psychological Review*, 65, 3, 151–166.

Newell, A., Shaw, J. C., and Simon, H. A. (1964). The process of creative thinking. In H. Gruber, G. Terell, and M. Wertheimer (Eds.), *Contemporary approaches to creative thinking*. New York: Atherton.

Newell, A., and Simon, H. A. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the ACM*, 19, 113–126.

Piccinini, G. (2010). The mind as neural software? Understanding functionalism, computationalism, and computational functionalism. *Philosophy and Phenomenological Research*, 86, 2, 269–311.

Pinker, S. (1997). *How the mind works*. New York: W. W. Norton.

Putnam, H. (1960). Minds and machines. S. Hook (ed.), *Dimensions of Mind* (New York: New York University Press), pp. 57–80.

Putnam, H. (1963). Brains and behavior. In R. Butler (ed.), *Analytical Philosophy Second Series* (Oxford: Basil Blackwell).

Putnam, Hilary. (1967). Psychological predicates. In W. H. Capitan and D. D. Merrill (eds.), *Art, Mind, and Religion* (Pittsburgh: University of Pittsburgh Press), pp. 37–48.

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Pylyshyn, Z. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, 4, 5, 197–207.

Rosenbloom, P., Laird, J., Newell, A., and McCarl, R. (1991). A Preliminary Analysis of

the Soar Architecture as a Basis for General Intelligence. *Artificial Intelligence*, 47,

289–325.

Rowlands, M. (1999). *The body in mind: Understanding cognitive processes.* Cambridge: Cambridge University Press.

Rupert, R. (1998). On the relationship between naturalistic semantics and individuation

criteria for terms in a language of thought. *Synthese*,117, 95–131.

Rupert, R. D. (2006). Review of *Embodiment and cognitive science* by Raymond Gibbs. *Notre Dame Philosophical Reviews*, 2006.08.20.

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.

Rupert, R. D. (2011). Embodiment, consciousness, and the massively representational

mind. *Philosophical Topics*, 39, 99–120.

Serino, A., and Haggard, P. (2010). Touch and the body. *Neuroscience and Biobehavioral Reviews*, 34, 224–36.

Shapiro, L. (2010). *Embodied Cognition*. New York: Routledge.

Simon, H. A. (1996). *Sciences of the Artificial,* Third Edition. Cambridge, MA: MIT Press. First edition published in 1969.

Simon, H. A., and Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist*, 26, 2, 145–159.

Sperry, R. W., Gazzaniga, M. S., and Bogen, J. E. (1969). Interhemispheric relationships: The neocortical commissures; syndromes of hemispheric disconnection. In P. J. Vinken and G. W. Bruyn (eds.), *Handbook of Clinical Neurology,* vol. 4 (Amsterdam: North-Holland), 273–290.

Steels, L., and Spranger, M. (2008). The robot in the mirror. *Connection Science*, 20, 4, 337–358

Stewart, J., Gapenne, O. & Di Paolo, E. (eds.) (2010). *Enaction*. Cambridge, MA: MIT Press.

Thelen, E., and Smith, L. (1994). *A dynamic systems approach to the development of*

*cognition and action.* Cambridge, MA: MIT Press.

van Gelder, T. (1995). What might cognition be, if not computation? *Journal of Philosophy* 92, 345–81.

Varela, F., Thompson, E., and Rosch, E. (1991). *The embodied mind.* Cambridge, MA: MIT Press.

Vera, A. H., and Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science* 17, 7–48.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of*  *Experimental Psychology,* 27, 635–657.

Weiskopf, D. (2004). The place of time in cognition. *British Journal for the Philosophy*

*of Science*, 55, 87–105.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, 8, 1, 44–57.

Wilson, R., and Foglia, L. (2011). Embodied cognition. *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition/> .

1. A few words about terminology are in order. Functionalism holds that the nature of a mental state is its causal role – its characteristic causal relations to inputs, outputs, and other mental states – not what kind of physical stuff its made of; functionalism thus allows that two creatures with different kinds of material bodies could be in the same mental state, so long as both creatures are in some state or other, of whatever composition, that plays the causal role individuative of the kind of mental state in question. Computational states are a kind of functional state, to be sure, individuated in terms of what they contribute to computational processing regardless of their physical composition. But, computational states are not necessarily a kind of mental state; so one could endorse a computational functionalism of sorts, without thinking computation has anything to do with the mind. In contrast, computationalism in cognitive science asserts that at some level of description (perhaps the neural level, perhaps the subpersonal level), a computational formalism provides the best way to model cognition-related processing or, ontologically speaking, that computational processing contributes significantly to the production of intelligent behavior (or other data of cognitive science). Computational functionalism results when a functionalist adopts an explicitly computational perspective about mental states mind, holding that mental states are, metaphysically speaking, functional states the nature of which is to play causal roles characteristic of computational states. For further discussion, see Piccinini (2010). [↑](#footnote-ref-1)
2. In the sense that at least some of the properties that play a causal-explanatory role are not identical to properties of independent interest in the physical sciences. They are, instead, multiply realizable, even if some of the properties – the embodiment-oriented ones – place stringent constraints on the range of physical structures that can realize them. [↑](#footnote-ref-2)
3. There are ways for embodiment theorists to resist. For instance, there’s an enormous literature on the so-called grounding problem (Harnad 1990), which suggests that for any mental representation to have content, it must be associated with sensorimotor states; an embodiment theorist who endorses this sensorimotor constraint on content *and* holds that the nature of sensorimotor states can’t be captured functionally will have genuinely set herself against functionalism. But, to make this anti-functionalist position credible, the embodiment theorist must give an alternative scientific account of the said sensorimotor states, and that has been thus far missing from the embodiment literature.

 Similar remarks apply to embodiment theorists who assign a privileged role to conscious experiences connected to the body or to certain biological processes. Consciousness and the maintenance of organismic integrity might play privileged roles in our understanding of cognition, but that does not itself speak against functionalism, unless accompanied by a scientifically respectable nonfunctionalist account of consciousness and of life. [↑](#footnote-ref-3)