



Book review

Michael Wheeler, *Reconstructing the Cognitive World: The Next Step*, MIT Press, June 2005. 6 x 9, 432 pp. \$35.00/£22.95 (CLOTH). ISBN-10: 0-262-23240-5; ISBN-13: 978-0-262-23240-1.

Strike while the iron is

Michael L. Anderson ^{a,b}

^a *Department of Psychology, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604-3003, USA*

^b *Institute for Advanced Computer Studies, Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD 20742, USA*

Available online 7 November 2006

Michael Wheeler opens his book with the following half-apology: “striking while the iron is lukewarm is sometimes the best strategy in philosophy” (Preface, p. x). He is referring to the fact that the main task of this book—to ferret out and critically assess the Cartesian legacy in contemporary cognitive science, replacing it (or at least altering it to better conform) with a philosophy of mind emphasizing the biological, embodied and embedded aspects of intelligence—was one he set for himself more than ten years ago. The work then culminated in his dissertation, completed in 1996. It is not a coincidence that this is the very same year my own dissertation was finished, on the very same subject. He is right that this topic was much on people’s minds at the time, and also right that it is less so now. The iron has indeed cooled.

The apology is misleading, however, in suggesting that the time for useful work on the topic has nearly passed. The fight over embodied cognition in the 1990s was less about forging philosophically sound foundations for a new kind of cognitive science than it was about creating institutional space to allow such work to occur. For this purpose the arguments for the new paradigm, and against the old, only had to be convincing; they didn’t have to be right. The current flourishing of embodied and situated approaches to AI, cognitive science and robotics has shown that the arguments from that period were indeed convincing to many, but time and reflection has in fact cast doubt on whether they were right. This is precisely the situation that most calls out for philosophical reflection. Indeed, if Socrates is any guide, the temperature of the iron is inversely proportional to the need (and, as he discovered in the most distressing way, the general desire) for disruptive reflections of the sort in which philosophers specialize. Philosophers strike while the iron *is*, and very often, annoyingly, while other people are using it.

As you will have figured out by now, *Reconstructing the Cognitive World* is a book by a philosopher and—although the author might wish otherwise—primarily for philosophers. There is much of interest here for the cognitive scientist, but any scientist who picks up this book must be willing not just to (re-)consider the foundations of her chosen field (a requirement that already narrows the candidate readers, albeit further than it probably should), but to do so by way of a heavy-duty, technical exegesis of a particular thread in the history of philosophy. The backbone of this work consists of a careful reading of Descartes’ philosophy of mind and nature, paired with an exegesis of Heidegger’s critique of, and alternative to, the Cartesian view. This is familiar territory to veterans of the fights between classical and embodied cognitive science, but, removed now from the battles of a decade ago, Wheeler cares less about whether Heidegger won, and more about whether he was right. Thus, Wheeler spends more than half of the book trying to figure out

E-mail address: anderson@cs.umd.edu (M.L. Anderson).

just exactly what Descartes said (not what you probably think), whether Heidegger's critique of Descartes is sound (not really), whether it can be fixed (not entirely) and what can be done about it (his implicit advice: if you are a philosopher, write a book like this one; if you are a cognitive scientist, let a philosopher into your lab so she can write a book like this one).

Wheeler's conclusion that Heidegger's arguments don't do the job they purport to do is clearly and convincingly supported. More of interest to the AI crowd will be his similar, and similarly well supported, conclusions that Hubert Dreyfus' arguments against computationalism don't succeed either, even when sympathetically reconstructed. This news is likely both to distress those scientists and philosophers who support the embodied-embedded account of cognition, and also to encourage some efforts to show that Wheeler is wrong to say that Dreyfus and Heidegger are wrong. But Wheeler doesn't say that they are wrong, only that their critiques of Cartesian-computational cognitive science are wrong. The question is what to do about it. Although clearly the available options include fixing the original critiques, or criticizing the critique critic (can one critique the critique critique?), another option is to deny that decisively undermining the Cartesian view is a necessary prerequisite for the construction of a viable alternative. Wheeler recommends, and he himself adopts, this last approach, thereby moving the philosophy of cognitive science away from critique and toward foundational exegesis. That is, he suggests that the more important philosophical task facing proponents of embodied-embedded cognition is to discover and make rationally consistent the assumptions that ground and motivate the actual practice of embodied-embedded cognitive science.

The idea that we should stop fighting about Descartes and just get on with the project of understanding, explaining, and (most importantly) *doing* embodied-embedded cognitive science seems sound enough, and one wonders if Wheeler might have taken the point more seriously himself, and dispensed with the long Cartesian preliminaries. Doing so might have made the book somewhat more attractive to cognitive scientists than it is likely to be in its current form, although it would also have made it less useful as a textbook-like guide to the *historical* origins of orthodox and embodied-embedded cognitive science. Possible use as a textbook may well be what the author—or the publisher—has in mind, but, although the book is well-written and thorough, it is pitched too high for the typical undergraduate classroom. Better, probably, to include the book on the recommended reading list for one's brightest students. In the end, although I recognize the scholarly-historical value of a scrupulous and detailed account of the post-Cartesian revolution, the fact remains that this part of the story has been told many times before. This may be the best account for the philosopher-scholar, but it is pretty clearly not the best for the student or the scientist.

This same criticism cannot be leveled against the material later in the book, where Wheeler gets serious about illuminating the conceptual foundations of embodied-embedded cognitive science (henceforth EECS). This material, while still difficult, should be of much wider interest and utility. Wheeler tackles this part of his project in three basic steps. First, he outlines what a Heideggerian cognitive science might look like, then he looks at the native conceptual foundations of embodied-embedded cognitive science, and finally he suggests that the Heideggerian perspective is the appropriate one to adopt in understanding the foundations of EECS.

It must be said that Wheeler does a good job here. His account of Heideggerian phenomenology and its implications for cognitive science is especially well done, and perhaps the best part of the book. According to Wheeler, the fundamental difference between Cartesian and Heideggerian cognitive science is to be found in their accounts of the content of basic experience. For Descartes, and for orthodox cognitive science (henceforth OCS), basic experience is composed of bits of context-independent sensory information—shapes, colors, motions—and the task of the thinking subject is to move from such context-independent features of experience to our more familiar sense of the world, which includes such elements as objects, people, animals, actions, situations, and the like. By contrast, in a Heideggerian cognitive science experience is always already a matter of encountering things in circumstances; one hears not “chug hiss rumble” but the train pulling into the station. Of course, this is not to say that the character of the sound *per se* is not a recoverable aspect of experience, but that for Heidegger it is the *decontextualization*, and not the contextualization that requires the conscious effort of the subject.

Now, this is a perfectly reasonable way to characterize a central difference between Descartes and Heidegger, and it is certainly useful as a backdrop to the further claim that for Heidegger (but not for Descartes) our fundamental encounter with the world is a matter of active, physical engagement in purposeful activity (and is not, for instance, primarily a matter of *observing* the world), but it might seem much less useful as a stand-alone distinction between OCS and EECS. The reason is that although many advocates of OCS might believe that we should account for the character of our experience in terms of mechanisms that take context-free input and manipulate it in various ways to produce our contextualized picture of the world, there are few if any who would argue that this is a largely conscious

process, or that the content of our basic conscious experience should be characterized in terms of shapes, colors, motions, and the like. Thus, it looks as if there is likely to be little if any disagreement between OCS and EECS with respect to the phenomenal character of experience. However, if we take the distinction to be a substantive claim not just about the phenomenal character of experience, but also about the mechanisms that produce it, then we can begin to get a better sense of what is at issue between Heideggerian and Cartesian cognitive science.

For instance, a distinctively Heideggerian cognitive science might be committed to an account of perceptual processing that is context-sensitive all the way down, from beginning to end. What might this mean? Well, rather than posit a series of largely automatic, encapsulated, unchangeable steps leading from, say, the impacts of photons on the retina to a 3-D picture of the visual space, a Heideggerian might expect a series of tunable filters, open to the influence of the interests, purposes and activities of the perceiver. And this is not a matter of influencing only what gets consciously noticed (although this is certainly crucial), but even influencing what impacts the retina. After all, the Heideggerian might point out, what one looks at—the pattern of saccades one makes—may well depend in part on what one is doing at the time.

Likewise, given the Heideggerian commitment to coping—active, purposeful, engagement with the world—as being our primary mode of encounter with the world, a Heideggerian cognitive science would be expected to emphasize the importance of know-how—the kinds of things we are able to *do* in the world—to structuring our experience of the world. Thus, we would be expected to be more attuned to affordances—relational properties specifying the appropriateness of things for certain uses or their availability for certain kinds of interaction—than to more “objective” properties like color, shape, and such.¹

This last point is reminiscent of Gibson (1979), and indeed, I have used Gibson’s term “affordances”, rather than the Heideggerian near equivalent “being-for”, but one of the services Wheeler provides to EECS here is to remind us that for Heidegger the mode of being where we are fully immersed in ongoing activity and primarily aware of objects as being-for our purposes (and aware of ourselves hardly at all, as we “lose” ourselves in acting), is only *one* of the possible modes by which we encounter the world. For Wheeler, one of the strengths of the Heideggerian account, and one of the reasons it is more appropriate for EECS than the Gibsonian view, is Heidegger’s analysis of two other modes of encounter with the world—roughly speaking, deliberative and reflective (ignoring the actual Heideggerian terminology)—and of the transitions and connections between them. Because Wheeler identifies the deliberative and reflective modes of encounter with the world as being the traditional subject matter of OCS, he suggests that an account of the human ability to transition between these various modes of encounter offers the eventual possibility of uniting OCS and EECS in one, comprehensive picture of human cognition.

This reconstruction of Heideggerian cognitive science is helpful and illuminating, as is his general account of the native intellectual foundations of EECS, but one place where I think Wheeler goes a bit astray is in his suggestion that the defining characteristic of EECS is a commitment to “non-trivial causal spread” in its account of cognition. The thought is this: EECS is typically committed to explaining cognition not in terms of the manipulation of stored internal representations, but rather in terms of agents interacting with their environments to solve various problems. For many proponents of EECS, the natural tool to use to explain such agents is not mathematical logic (e.g. to describe the rules of thought by which representations are transformed), but dynamical systems theory, which can be used to describe the trajectory of a complex system through an abstract state space. Because the state of an agent at time $t + 1$ naturally depends not just on its internal state at time t , but also on the state at time t of those entities with which it interacts, the equations defining the state transitions of the agent will have terms representing the state(s) of those entities. And if, as one might expect, the states of those entities at time t depends not just on their internal states at time $t - 1$, but also on (some of the) states and actions of the agent at $t - 1$ (or $t - n$), then the equations describing the state transitions of the entities will require terms for the agent. The result is a set of highly coupled equations in which one cannot account for the behavior of the agent except by reference to (the states and contributions of) entities

¹ Note that I have exaggerated the differences between the positions for the sake of pedagogical clarity; it is unlikely that any advocate of EECS, however influenced by Heidegger, would deny the existence of some context-free processing mechanisms, nor would any defender of OCS, however big a fan of Descartes, deny that some lower-level cognitive/perceptual processes may be open to the influence of, e.g., beliefs or attention, or that know-how might play some role in cognition. In the wild, the distinction between OCS and EECS will manifest itself not in terms of stark oppositions, but in differing predictions about whether context-free mechanisms and objective, propositionally structured knowledge, or rather context-dependent mechanisms and non-conceptual know-how, will turn out to underpin the bulk of cognitive processing.

outside of the boundaries of the agent. Indeed, many theorists of EECS suggest that one typically ought to talk about the agent and the various implicated entities as different parts of a single coherent cognitive system.

This last step is the essence of “non-trivial causal spread”. Suppose one asks the simple question of why (perceptual) input X led to behavior Y for a given agent. If the cognitive system carrying out the input-output transformation exhibits non-trivial causal spread, then one cannot answer this question solely by reference to the inner states and mechanisms of the agent, but must implicate elements of the environment as well. Such spread of relevant causes to elements of the environment is “non-trivial” when the elements implicated are not mere background conditions (for Y to happen one condition was that the Earth remained solid), but active contributors to the specific outcome (part of the process resulting in Y included agent-caused alterations of the environment without which the result would not have been possible). The classic example is long division. When asked “What is 39465 divided by 5?” (input X), the agent says “7893” (behavior Y) because she was able, as part of the cognitive processing, to write down intermediate results of the long division on a piece of paper.

It is certainly true that a great deal of work in EECS emphasizes the ways in which behavior is a fundamental part, and not just the outcome, of cognitive processes. And Wheeler is quite right to notice that when actions are integrated into cognitive processes, or when cognitive aids or technologies are employed as part of a problem-solving routine, then it can be illuminating to view the cognitive system as extended to include those elements of the body and world making non-trivial causal contributions to the cognitive outcome. But there is a great deal of other work squarely in the EECS camp for which it is much less clear that the notion of non-trivial causal spread offers much explanatory value. For instance, work on the role that neural resources involved in motor-control might play in supporting higher-order cognitive processes like language understanding [1,3], and investigations into the influence that metaphorical mappings from elements of our embodied experience like moving around or standing upright might have on the way we think about abstract things like planning or morality [2], clearly fall under the rubric of EECS, yet do not seem to require any reference to non-trivial causal spread in their explanations of the phenomena under consideration. The problem is not that Wheeler makes the outrageous claim that all legitimate explanations in cognitive science must involve non-trivial causal spread; rather, he argues that a defining characteristic of EECS is the belief that all intelligence ultimately grounds out in interactive coping (“on-line intelligence”) of the sort that exhibits non-trivial causal spread. The theorists mentioned above may endorse that view, but their positions (and those of several others) are also compatible with the less radical thought that the neural circuits that explain intelligent behavior were shaped by such interactive coping in our evolutionary past, and still bear the marks of that origin, which fact is naturally important to a full understanding of the nature and operations of those circuits. It doesn’t seem to me that such conservatism should be grounds for booting someone out from under the EECS umbrella.

This would be just a quibble, except that it points to a somewhat larger problem for the book as a whole. While the individual elements of Wheeler’s explanations are sound and helpful—from his reconstruction of the Heideggerian position on mind, to his accounts of dynamical systems theory—it is not clear they add up to the unifying account of EECS that Wheeler is aiming for. It may be that Wheeler is a victim of his own clarity, but his accounts of the native conceptual resources of EECS (dynamical systems theory, emergent behavior, cognitive actions, non-trivial causal spread, etc.) leave one with the impression that the field is already doing quite well explaining itself to itself, without the need for any Heideggerian re-structuring. Wheeler certainly *does* make the case for the compatibility of EECS and Heideggerian phenomenology, and his observations may therefore help other Heideggerians to understand EECS, and to see how it fits into a broader picture of human life, but what never becomes clear is what it is about EECS—what about it is unexplained or unexplainable with existing resources—that would motivate anyone to adopt the Heideggerian world-view.

This failure appear more or less serious depending on what one hopes to get from the book. As a high-level, scholarly, and thorough introduction to the independently interesting and historically related subjects of Heideggerian phenomenology and embodied-embedded cognitive science, this book is an enjoyable success. But as an attempt at a novel re-construction of the foundations of EECS likely to influence the way that discipline is understood by its practitioners, *Reconstructing the Cognitive World* falls somewhat short of the mark. As an EE cognitive scientist myself, I must admit that I find it a bit disappointing, to agree (by reading such a book) to subject one’s disciplinary tools—however imperfectly forged—to the heat and hammering of philosophical analysis, only to have them left unchanged, apparently just fine the way they were. Since EECS, as with any scientific discipline, can benefit immensely from a fruitful re-working of its founding assumptions, one is left to hope the next such strike is somewhat more transformative.

References

- [1] A. Glenberg, M. Kaschak, Grounding language in action, *Psychonomic Bulletin and Review* 9 (2002) 558–565.
- [2] G. Lakoff, M. Johnson, *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books, 1999.
- [3] F. Pulvermüller, Brain mechanisms linking language and action, *Nature Reviews Neuroscience* 6 (2005) 576–582.