

\*\* Draft Version. The final version is forthcoming in *Minds and Machines*.

**Robert D. Rupert**, *Cognitive Systems and the Extended Mind*, Oxford University Press, 2009, xiii+268, \$55, ISBN 9780195379457

1. Suppose there's this establishment. It's been around for a while; it's served its purpose; it's paid its dues. Although there might have been minor troubles here and there, dissident voices now and then, nothing was too serious; all was well. Recently, however, events have transpired and a change of mind in some has occurred: now, it's said, the establishment is chauvinistic, overly conservative, straightforwardly narrow-minded. It's time for a revolution, these some assert. A more liberal, more progressive, more exciting enterprise is in order.

Vague as it is, the above description captures at least the general atmosphere of the current debate concerning the location of the vehicles of cognition. The establishment, *vis-à-vis*, the traditional, computationally inspired, and organismically bounded approach to human cognition, has come under fire. Recent (and not so recent) philosophical and scientific findings suggest a picture of human cognition that diverges significantly from that delineated by the orthodoxy. Ask 'Where's the mind?' or, better, ask 'Where are the vehicles of cognition?,' and two mutually exclusive answers will be given. Friends of the establishment locate cognition solely within the agent's skin or skull; foes of the establishment locate cognition only partly within these aforementioned, ancient boundaries. This latter view, the *extended* view, holds that the vehicles of cognition transcend the organism and extend into the environment. Natural, technological, and even social resources can all be literal parts of the mind.

Assume that some human cognition is *actually* extended. What follows? A whole lot. Cognitive neuroscience is in trouble; computational psychology is partly misguided; and social psychology is in need of a change. Not even our common conception of mentality is safe. Expressions such as 'I see your desire (and I raise it)' or 'You are sitting on my memory!' will (finally) make sense and cease to be taken as signs of dementia. Extensionphiles don't only promise change, but liberation too: 'once the hegemony of skin and skull is usurped, we may be able to see ourselves more truly as creatures of the world' (Clark and Chalmers 1998, p. 18).

How do you stop this revolution from happening? Here's one way: show that whatever has been discovered or transpired has been oversold and doesn't

undermine the establishment. In fact, show that the extant resources of the establishment—perhaps, with some minor adjustments—adequately do the job. Furthermore, show that this new, more liberal, more exciting enterprise doesn't have much to offer. *Cognitive Systems and the Extended Mind* is an attempt to show all that, and more. Its conclusion: there's nothing about the establishment which is chauvinist, overly conservative, or narrow-minded. Instead, the establishment privileges only that which deserves to be privileged and is suitably realistic. Above all, the establishment is rightly—to the right extent, that is—mindful.

If some books are revolutionary, *Cognitive Systems* isn't. It's a remarkable book nonetheless. Without a doubt, it makes a great contribution to the field.

2. *Cognitive Systems* is comprised of three parts, the first two of which are a direct attack on the extended view. Taken together, they argue for the following claims: (i) There's a competing view to the extended view; (ii) There are no reasons to prefer the extended view over its non-extended competitor; and (iii) There are reasons to reject the extended view. Let me elaborate.

The extended view isn't the only player in town. There's an alternative and competing view, the *embedded* view, according to which, 'typical cognitive processes depend, in surprising and complex ways on the organism's use of resources . . . but cognition does not literally extend into the environment' (p. 5). The embedded view, Rupert tells us, has its virtues. It accommodates for the empirical findings that impress extensionphiles and naturally accounts for the successes of cognitive psychology. What's more, it's a simpler and more conservative view: unlike the extended view, it neither postulates new entities nor does it re-label cognitive processes. Unless there are substantial reasons that favor the extended view, there's little or no motivation to choose it over the embedded.

But such reasons are nowhere to be found, Rupert argues. Neither analytical functionalism nor psychofunctionalism do much to promote the extended view. Functionalist considerations either fail to yield extended kinds or, in case they do so, they render these kinds explanatorily useless (Chapter 5). Arguments based on the transformative powers of extended resources don't yield extended systems. The general form of these arguments is found to be problematic, and specific considerations regarding the role of language in cognition do little to ameliorate the situation (Chapter 6). The same verdict

is reached for arguments which conclude that what ‘contributes causally, in a distinctive and nontrivial way, to the production of some cognitive phenomenon’ is itself cognitive (p. 19). Introducing considerations from either developmental systems theory or dynamical systems theory, in the hope of showing ‘that cognitive science comprises a special domain for which a circumscribed form’ of the argument holds, still makes no difference (p. 175). The first move falters, since the components of the supposed extended system are reproductively independent and as such don’t form an individual (Chapter 5). The second move fails, for although dynamical-systems-based models of cognition could, in principle, support the extended view, extant dynamical-systems-based models don’t (Chapter 7). Finally, Rupert shows that arguments for the extended view that stem from the contents of conscious experiences are (predictably) weak (Chapter 8).

And if this isn’t enough, Rupert offers two reasons to reject the extended view. First, the only plausible criterion of demarcation between what’s cognitive and what’s not locates cognition exclusively inside the boundaries of the organism (Chapter 3). Second, the extended view runs into trouble when it comes to explaining the practice of cognitive psychology. It either ‘makes little sense’ of such a practice, or, it accounts for it, but in doing so offers no advantage over the embedded view (p. 46).

In the third and final part, Rupert turns away from the extended view. He instead argues that findings from the embedded or situated programs in cognitive science are consistent with the orthodox, computationally inspired picture of the human mind. These findings neither suggest an anti-computationalist outlook (Chapter 9), nor do they require a form of representation, which is in principle unavailable to the computational approach (Chapter 10). Finally, in the last substantial chapter of the book, Rupert argues that no plausible version of the *embodied* view—namely, the view that human cognition depends on, or is determined by, our bodily constitution—is at odds with the orthodoxy. The establishment is safe, Rupert shows. If anything, the establishment now looks even stronger.

3. I don’t have much to say about the third part of the book, nor about Rupert’s conservative (deflationary, if you prefer) reading of the empirical support for the extended view. My focus will be elsewhere: first, on his critique of functionalist arguments in support of the extended view; and, second and only briefly, on his contention that the extended view fails to accommodate for the success of cognitive psychology.

Functionalism is the theory of choice for the extensionphiles, and for good reason: assuming functionalism, the possibility of extended cognition is established. If all there is to being a cognitive state is fulfilling a certain role, then the location of the vehicles of cognition comes out to be, at least on a theoretical level, irrelevant. But the possibility of extended cognition is not enough. To get the extended view, one also needs to add that there are actual realizers of cognitive states with parts that extend beyond the organism. Here, however, is where trouble starts for the proponent of the extended view. Or so, Rupert claims.

For argument's sake, suppose that analytical functionalism doesn't yield extended cognitive processes. The proponent of the extended view, if she doesn't have a criterion of demarcation of her own, needs to turn to cognitive science for guidance. How else will she know what's cognitive? But cognitive science, Rupert argues, provides fine-grained functionalist profiles, which speak against the extended view: purported examples of extended states turn out to be non-cognitive, for they don't exhibit the requisite (fine-grained) similarity to paradigmatic inner psychological states. Faced with this difficulty, the proponent of the extended view can propose to individuate mental states according to coarse-grained causal-functional profiles. This taxonomical arrangement gives rise to extended kinds. What it fails to do, Rupert tells us, is to offer a causal-explanatory advantage. Individuating mental kinds broadly remains unmotivated.

Clark has recently taken up Rupert's challenge. He offers the following two-fold rejoinder. First, he points out that 'acceptable forms of unification need not require all systemic elements to behave according to the same laws' (Clark (2008), p. 115). In fact, to require this, 'is simply to beg the question against any science whose target is a genuinely hybrid system' (ibid.). We should instead leave this issue up to science: 'it is the substantive empirical bet of the extended systems theorist that the larger hybrid wholes, comprising biological and nonbiological elements, will [sic] *also* . . . prove to be the proper objects of sustained scientific study in their own right' (ibid.). This, however, fails to assuage Rupert's concern. Clark doesn't provide evidence that bio- and extended-memory, for instance, form a uniform kind. He merely offers a promissory note. Yet, stating that P could be wrong, doesn't amount to much of an argument against P. No one holds that, *necessarily*, cognition doesn't extend. Critics of the extended view are only committed to the view that, *given* current evidence, we should reject the extended view.

To be fair, Clark acknowledges that this reply doesn't get to the heart

of the matter, hence, the second part of his reply: ‘[e]ven if Rupert and others are right that terms such as *memory* cannot, once extended to the nonbiological domain, themselves pick out explanatorily unified kinds, this does not mean that the extended organizations in which they participate are not proper objects of scientific enquiry’ (Clark (2008), p. 121). And, how, do we know that they can be the objects of scientific enquiry? We know this because there is ‘[already] a nascent science both of the recruitment (of sets of neural and extraneural resources) and of the fine-tuned unfolding of activity in just such heterogeneous ensembles’ (ibid.). But this response still falls short, for Rupert makes a plausible case that the findings of the nascent science that Clark has in mind are ambiguous between an extended and an embedded reading (see pp. 102–5, for Rupert’s embedded-friendly reading of the findings). To provide support for the extended view, Clark needs to show that the embedded view fails to account for these findings. Until then, Rupert’s challenge stands.

Wheeler (forthcoming) offers a different response to Rupert’s concern regarding the costs of coarse-grained individuation. ‘Imagine,’ Wheeler says, ‘we came across a human being whose purely inner memory system didn’t exhibit the generation effect, but who nevertheless continued to achieve the context-sensitive storage and retrieval of information’ (xx). Wheeler contends that no cognitive psychologist would conclude that the ‘subject lacks the cognitive trait of memory’ (xx). Most importantly, Wheeler adds, the ‘subject’s abilities would undoubtedly be investigated by the cognitive psychologists as one possible form of the psychological phenomenon of memory’ (xx). These two claims license Wheeler to conclude that ‘the generic notion of memory that underwrites this way of proceeding is doing important work in organizing and shaping the project of cognitive-scientific explanation’ (xx). Extended-cognitive science is vindicated after all.

Wheeler’s argument goes as follows: there’s a psychological kind X, which is individuated, cognitive science tells us, by a fine-grained causal-functional profile. For instance, a psychological state M is X, if M is whatever does X1, X2, X3, and X4. Now, suppose there’s a psychological state, Y, which does X1, but not X2, X3, and X4, and its realizer is located inside the organism. Suppose further that science takes an interest in Y and judges it to be a member of Z, a kind which subsumes both X and Y. Finally, the argument concludes, if there’s another state, Y\*, which exhibits X1 (but not X2, X3, and X4) but its realizers are located (at least, partly) outside the organism, then it would be question begging to hold that science would be interested

in  $Y$  but not in  $Y^*$ . Hence, the argument's crucial premises: (P1) Science is interested in  $Y$ ; and (P2) It's question begging to accept (P1) for  $Y$ , but deny it for  $Y^*$ .

The first thing to note is that folk ascriptions of mentality have little or no bearing on this issue. We should look elsewhere for reasons to accept (P1). Perhaps, considerations from the practice of cognitive science ground (P1). For instance, one may argue that cognitive science is (or should be) interested in  $Y$  because  $Y$  is a case of  $Z$ . In terms of Wheeler's example, cognitive science takes an interest in inner information storage and retrieval because it judges it to be a case of memory (where memory is treated as a broad kind). But to get from 'Y is a case of Z' to 'Cognitive science is interested in Y' we need to add two premises: first, that  $Z$  itself is a cognitive/psychological kind; and second, if  $Z$  is of interest to cognitive science, then whatever is a case of  $Z$  is also of interest to cognitive science.

However, if Rupert (and others) are right to insist that 'the most interesting and useful profiles of psychological states and properties detail causal roles that external materials are not likely to fill,' then  $Z$  is unlikely to be a cognitive/psychological kind. Consequently, the *current* practice of cognitive science doesn't ground (P1) (pp. 95-6). Is this response question begging? Have I assumed that psychological states need to be individuated finely, even though we are now considering the possibility of coarse-individuation? Not exactly. The claim is that our best bet for figuring out what's cognitive is to look to the practice of cognitive science. We need to start from somewhere and cognitive science seems the obvious place. Cognitive science's practice of individuation speaks against (P1). Contingently then, we have reasons to reject (P1).

In response, one can either show that current cognitive science isn't solely concerned with fine-grained causal-functional roles, or deny that the current object of study of cognitive science should determine its future object of study. Consider the first possibility. The aim is to show that, despite what came before,  $Y$  is a cognitive/psychological kind. One can motivate this move by insisting that Wheeler's example does indeed show that cognitive science is interested in  $Y$ . And if cognitive science is interested in  $Y$ , then  $Y$  is a cognitive/psychological kind. But the argument, 'Cognitive science is interested in  $Y$ ; whatever cognitive science is interested in is a cognitive/psychological kind; therefore,  $Y$  is a cognitive/psychological kind,' isn't a very good one. It can't be the case that whatever cognitive science is interested in, this turns out to be a cognitive kind. Additional requirements need to be speci-

fied. But here a tension arises. If the term ‘taking an interest in something’ is understood demandingly—for instance, cognitive science is interested in Y when the involvement of Y in the theoretical framework of cognitive science results in an increase in explanatory power—then, although the premise ‘whatever cognitive science is interested in is a cognitive/psychological kind’ is rendered true, it’s doubtful that cognitive science is interested (in this demanding sense) in Y. Conversely, if ‘taking an interest in something’ is understood loosely, that is, in a manner that makes it true that cognitive science is interested in Y, then the premise ‘whatever cognitive science is interested in is a cognitive/psychological kind’ will most likely be false.

Now consider the latter possibility, namely, the contention that the current object of study of cognitive science shouldn’t determine its future object of study. The idea here seems to be the following: suppose we have a criterion of what counts as cognitive, which isn’t wholly determined by current cognitive practice. If current cognitive practice says that Y isn’t of interest, since it belongs to a too broad of a kind to give rise to interesting generalizations, yet our criterion yields that Y is cognitive, we should accept the verdict of the latter. Consequently, we should divert our scientific practice in such a way as to study Y. If this proposal is accepted, then it is clear that Wheeler’s argument depends on the merits of the account (criterion) of cognition that we provide in its support. Whatever the specifics of this account might be, it must be broad enough to include Y. Most importantly, it should independently motivate the inclusion of Y in the set of the objects of study of cognitive science. That is to say, since the verdict of this account diverges from that of current cognitive science, one needs to show that there’s an explanatory benefit from adopting this account. But this takes us back to where we started: Rupert’s demand for proof that coarse individuation is explanatorily beneficial.

4. All other things being equal, Rupert argues, conservatism and simplicity favor the embedded view. A choice of the extended view remains unmotivated, for it fails to deliver substantial benefits to counterbalance the loss of simplicity that results from the introduction of new categories. Rupert provides another, more direct, argument against the extended view. According to Rupert, ‘[a] significant amount of successful research in cognitive psychology . . . presupposes the existence of human cognitive systems of temporal grain  $\delta t$ ’ (p. 59). Yet a commonly (perhaps, the most commonly) used principle of demarcation (the causal principle), which does deliver extended cognitive systems, yields a temporal grain smaller than  $\delta t$ . Therefore,

the principle must be rejected for it ‘compromises the research successes’ of cognitive psychology (ibid.).

There’s much to be said about this argument. For instance, is cognitive science monolithic in this (temporal) respect? Does it need to be? Furthermore, is there one appropriate temporal grain for all cognitive systems, or does the temporal grain of each system need to be established case-by-case? (see also Wilson (2010)). These points need to be addressed, if Rupert’s argument is to succeed. Be that as it may, if Rupert is correct to point out that the duration of a system matters for whether the system is cognitive, then although functionalism might have liberated cognition from its material underpinnings and (at least, theoretically) from the location of its realizers, it hasn’t succeeded in liberating cognition from temporal constraints. In our attempt to better understand the nature of cognition we should, perhaps, not only ask ‘*What* is a mind?’ or ‘*Where* is a mind?’ We should also ask ‘*When* is a mind?’

5. There’s much to be admired in this book, and not much left to be desired. Without sacrificing philosophical rigor or attention to empirical details, Rupert repeatedly brings empirical findings under philosophical scrutiny. From start to finish, he spots enthymematic arguments, and over and over again, he challenges both opponents of the embedded view and those who hold that the rules and representations approach to cognition is outdated and needs to be supplanted. *Cognitive Systems* demands the attention of everyone who is interested in the nature of cognition. I highly recommend this book.

\*\* I would like to thank Michael Wheeler for his helpful comments.

## References

- Clark, A. (2008). *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.
- Clark, A., and Chalmers, D (1998), The Extended Mind. *Analysis* 58, 7-19.
- Wheeler, M. (forthcoming), ‘In Defence of Wide Functionalism,’ in Richard Menary (ed.), *The Extended Mind*. Cambridge, MA: MIT Press.
- Wilson, R. (2010) Review of *Cognitive Systems and the Extended Mind* by Robert D. Rupert, in Notre Dame Philosophical Reviews, 2010.03.07.

Andreas Elpidorou  
Department of Philosophy

Boston University  
745 Commonwealth Ave.  
Boston, MA 02215, U.S.A.  
E-mail: [aselpido@bu.edu](mailto:aselpido@bu.edu)